ADILSON DA SILVEIRA

LAGRANGEANO AUMENTADO EXPONENCIAL APLICADO AO PROBLEMA DE EQUILÍBRIO

Atena Editora
Ano 2023

ADILSON DA SILVEIRA

LAGRANGEANO AUMENTADO EXPONENCIAL APLICADO AO PROBLEMA DE EQUILÍBRIO

Atena Editora

Ano 2023

CIÊNCIAS EXATAS E DA TERRA E ENGENHARIAS

Lagrangeano aumentado exponencial aplicado ao problema de equilíbrio

Diagramação: Ellen Andressa Kubisty
Correção: Maiara Ferreira
Indexação: Amanda Kelly da Costa Veiga
Revisão: O autor
Autor: Adilson da Silveira

Dados Internacionais de Catalogação na Publicação (CIP)

S587 Silveira, Adilson da
Lagrangeano aumentado exponencial aplicado ao problema de equilíbrio / Adilson da Silveira. – Ponta Grossa - PR: Atena, 2023.

Formato: PDF
Requisitos de sistema: Adobe Acrobat Reader
Modo de acesso: World Wide Web
Inclui bibliografia
ISBN 978-65-258-1681-4
DOI: https://doi.org/10.22533/at.ed.814231107

1. Otimização matemática. 2. Programação matemática. 3. Problemas de equilíbrio. 4. Método do Lagrangeano aumentado. I. Silveira, Adilson da. II. Título.
CDD 519.6

Elaborado por Bibliotecária Janaina Ramos – CRB-8/9166

Atena Editora
Ponta Grossa – Paraná – Brasil
Telefone: +55 (42) 3323-5493
www.atenaeditora.com.br
contato@atenaeditora.com.br

DECLARAÇÃO DO AUTOR

O autor desta obra: 1. Atesta não possuir qualquer interesse comercial que constitua um conflito de interesses em relação ao conteúdo publicado; 2. Declara que participou ativamente da construção dos respectivos manuscritos, preferencialmente na: a) Concepção do estudo, e/ou aquisição de dados, e/ou análise e interpretação de dados; b) Elaboração do artigo ou revisão com vistas a tornar o material intelectualmente relevante; c) Aprovação final do manuscrito para submissão.; 3. Certifica que o texto publicado está completamente isento de dados e/ou resultados fraudulentos; 4. Confirma a citação e a referência correta de todos os dados e de interpretações de dados de outras pesquisas; 5. Reconhece ter informado todas as fontes de financiamento recebidas para a consecução da pesquisa; 6. Autoriza a edição da obra, que incluem os registros de ficha catalográfica, ISBN, DOI e demais indexadores, projeto visual e criação de capa, diagramação de miolo, assim como lançamento e divulgação da mesma conforme critérios da Atena Editora.

CIÊNCIAS EXATAS E DA TERRA E ENGENHARIAS

DECLARAÇÃO DA EDITORA

Este trabalho é dedicado a minha esposa Beatriz e
às minhas filhas Roberta e Luiza.

AGRADECIMENTOS

A Deus pela proteção na caminhada e pela graça de concluir este trabalho.

Ao meu orientador, professor Doutor Luiz Carlos Matioli, pelos ensinamentos, moti- vação, persistência, incentivo e paciência, sempre respeitando e compreendendo minhas dificuldades e limitações.

À Universidade Federal do Paraná e ao Programa de Pós-Graduação em Métodos Numéricos em Engenharia, pela oportunidade de cursar o Doutorado.

Aos funcionários e grupo docente do Programa de Pós-Graduação em Métodos Numéricos em Engenharia, pelos valiosos ensinamentos transmitidos.

À Universidade Tecnológica Federal do Paraná, Câmpus Pato Branco, por ter me proporcionado a oportunidade de realizar esse curso.

O presente texto traz os resultados mais importantes obtidos a partir do trabalho de pesquisa desenvolvido pelo autor, durante a elaboração da sua tese de doutorado, junto ao Centro de Estudos de Engenharia Civil – CESEC, da Universidade Federal do Paraná, junto ao programa de Pós-Graduação em Métodos Numéricos em Engenharia.

O tema principal desta obra versa sobre métodos de lagrangeano aumentado, os quais são conhecidos, populares e estudados pela comunidade acadêmica que trabalha na área de análise numérica e otimização. Boa parte desta popularidade do método, se deve a sua grande aplicabilidade na resolução de problemas práticos ou teóricos, principalmente nas áreas de engenharia e na matemática aplicada. Tais métodos têm em sua estrutura, como componente principal, uma função (termo) de penalização. As diferentes variações dessa função de penalização podem dar origem a diferentes métodos de lagrangeano aumentado. Seguindo esta metodologia, neste trabalho apresenta-se um novo método de lagrangeano aumento cuja função de penalização é do tipo exponencial. Apresenta-se o modelo, algoritmos relacionados, resultados sobre condições de convergência e implementações. Por fim, esse método é aplicado na resolução de uma classe específica de problemas matemáticos importantes: problemas de equilíbrio. Nesse contexto são apresentados resultados numéricos referentes às implementações dos algoritmos, análises, conclusões sobre os resultados e, também, apontados caminhos para a continuidade da pesquisa.

LISTA DE ABREVIATURAS E SIGLAS

EP	Problema de Equilíbrio
IALEM	Método de Lagrangeano Aumentado Extragradiente Inexato
IPPEM	Método de Ponto Proximal Extragradiente Inexato
IPPEMI	Método de Ponto Proximal Extragradiente Inexato Irrestrito
IALEMe	Método de Lagrangeano Aumentado Extragradiente Exponencial Ine-xato
IPPEMIe	Método de Ponto Proximal Extragradiente Inexato Irrestrito
NEP	Problema de Equilíbrio de Nash
GNEP	Problema de Equilíbrio de Nash Generalizado

SUMÁRIO

RESUMO

Neste estudo propomos um Método de Lagrangeano Aumentado Exponencial para resolução do Problema de Equilíbrio Geral. Tal método é uma extensão dos algoritmos apresentados em (NASRI, 2010). Em nossa proposta, substituímos a penalidade clássica de Rockafellar presente nestes algoritmos pela penalidade exponencial. Em seguida, refizemos a teoria geral em torno do novo algoritmo. Mostramos que o mesmo está bem definido, que possui terminação finita, que admitidas às condições iniciais adequadas o mesmo gera uma sequência convergente. A teoria de convergência é construída via ponto proximal, ou seja, a partir da relação de dualidade entre lagrangeano aumentado e ponto proximal, porém o termo de regularização aqui utilizado é uma quase distância de Bregman, diferente do termo quadrático empregado em (NASRI, 2010). Para contornar possíveis problemas de mau condicionamento causados pela presença da penalidade exponencial, procedemos a um ajuste quadrático da mesma. Em seguida, testamos a nova metodologia por meio de experimentos numéricos, considerando inicialmente o método com a penalidade exponencial e, em seguida, com a quadrática ajustada. Para realizar os testes, escolhemos problemas de equilíbrio e problemas de equilíbrio de Nash generalizados (GNEPs), os quais compõem uma classe particular de problemas de equilíbrio. O método puro resolveu problemas pequenos e, para problemas com dimensões maiores, a versão com quadrática ajustada mostrou-se melhor. Na sequência, discutimos os resultados numéricos e apresentamos as considerações finais. Para finalizar, deixamos algumas perspectivas de trabalhos futuros e uma lista de referências, as quais serviram de suporte para essa pesquisa.

PALAVRAS-CHAVES: Palavras-chaves: Problema de Equilíbrio, Lagrangeano Aumentado, Penalidade Exponencial.

ABSTRACT

In this study we propose an Exponential Augmented Lagrangian Method to the resolution of the General Equilibrium Problem. Such a method is an extension of the algorithms presented in (NASRI, 2010). In our proposal, we replaced the classic penalty Rockafellar present in these algorithms by the exponential penalty. Then, we redid the general theory surrounding the new algorithm. We show that it is well defined, has finite termination and, admitted the appropriate initial conditions, it generates a convergent sequence. The convergence theory is built via proximal point, ie, from the dual relationship between Augmented Lagrangian and the proximal point, but the regularization term used here is almost distance Bregman, different from the quadratic term used in (NASRI, 2010). To work around possible bad conditioning problems caused by the presence of the exponential penalty, we proceeded to a quadratic adjustment of it. Next, we tested the new method by numerical experiments, considering initially the method with the exponential penalty and then with the quadratic adjusted. To perform the tests, we chose equilibrium problems and Generalized Nash equilibrium problems (GNEPs), which compose a particular class of equilibrium problems. The pure method solved little problems and for problems with larger dimensions, the version with the quadratic adjusted proved to be better. Following, we discuss the numerical results and present the final considerations. Finally, we leave some perspectives of future works and a list of references, which served as support to this research.

KEY-WORDS: Equilibrium Problems, Augmented Lagrangian, Exponential Penalty.

INTRODUÇÃO

1 I MOTIVACÃO

Em anos recentes o problema de equilíbrio tem recebido atenção dos pesquisadores. Diversos resultados teóricos e experimentos numéricos têm sido publicados. Entre os trabalhos desta natureza podemos apontar (BIANCHI; PINI, 2005), (BIGI et al., 2013) e (BLUM; OETTLI, 1994). Ainda, sobre esse tema destacamos (BURACHIK, 2012), (FACCHINEI, 2003), (IUSEM; NASRI, 2007), (IUSEM; KASSAY; SOSA, 2009), (IUSEM; SOSA, 2003a), (IUSEM; SOSA, 2003b), (MASTROENI, 2000; MATIOLI; SOSA; YUAN, 2012; MUU; OETTLI, 1992; MUU; QUOC, 2009; FACCHINEI; KANZOW, 2007; QUOC; MUU, 2004; TRAN; DUNG; NGUYEN, 2008; SANTOS; SCHEIMBERG, 2011; NGUYEN; STRODIOT; NGUYEN, 2009; NGUYEN; STRODIOT et al., 2009). Em particular (BIGI et al., 2013) apresenta um apanhado sobre algoritmos e condições de existência de soluções para esta classe de problemas.

Como casos particulares do problema de equilíbrio destacamos: problemas de otimização convexa, problemas de ponto fixo, problemas de complementaridade, problemas de equilíbrio de Nash, problemas de desigualdade variacional e problemas de minimização vetorial, como destacado nos trabalhos de (BIGI et al., 2013; BLUM; OETTLI, 1994; IUSEM; SOSA, 2003a).

Em (NASRI, 2010) são apresentados métodos de lagrangeano aumentado para resolução de problemas de equilíbrio em dimensão finita, cujos conjuntos viáveis são definidos por inequações convexas, generalizando o método de lagrangeano aumentado proximal para otimização restrita. A cada iteração as variáveis primais são atualizadas resolvendo-se um problema de equilíbrio irrestrito e, então, as variáveis duais são atualizadas por uma fórmula fechada e a análise de convergência é apresentada.

Na sequência definimos o problema de equilíbrio com as propriedades mínimas necessárias para o desenvolvimento teórico.

Definição 1.1 *Considere K um subconjunto de R^n não vazio, convexo e fechado e uma função $f : K \times K \to R$ satisfazendo*

P1: $f(x, x) = 0$ para todo $x \in K$,

P2: $f(x, \cdot) : K \to R$ é convexa e semicontínua inferior para todo $x \in K$,

P3: $f(\cdot, y) : K \to R$ é semicontínua superior para todo $y \in K$.

O problema de equilíbrio, denotado por $EP(f, K)$, consiste em determinar $x^* \in K$ tal que $f(x^*, y) \geq 0$ para todo $y \in K$. O conjunto de soluções para $EP(f, K)$ será denotado por $S(f, K)$.

Em termos de métodos computacionais para problemas de equilíbrio destacamos

os trabalhos de (BIGI et al., 2013; FACCHINEI; KANZOW, 2010a). Em (NASRI, 2010), também são introduzidas versões exatas e inexatas de métodos de lagrangeano aumentado para resolução do EP (f, K) em R^n, para o caso nos quais o conjunto viável K é da forma:

$$K = \{x \in R^n \ : h_i(x) \leq 0 \quad (1 \leq i \leq m)\} \tag{1.1}$$

em que $h_i : R^n \rightarrow R$ são funções convexas para todo $i = 1, \ldots, m$.

Observe que na Definição 1.1, há a necessidade das duas variáveis, x e y, estar no conjunto convexo K, ou seja, resolver o EP (f, K) consiste em determinar x* $\in K$ tal que f $(x^*, y) \geq 0$, para todo y em K.

Devido ao fato das variáveis pertencerem ao conjunto convexo K, a maioria dos artigos encontrados na literatura recorrem a algum tipo de projeção para atender essa condição. É bem conhecido que fazer projeção resulta em outro problema de otimização que pode ser um problema de difícil resolução, a menos que o conjunto K seja tipo poliedros.

No caso dos métodos de lagrangeano aumentado, a "dependência de estar no conjunto" será enfraquecida através da abordagem de penalização.

2 | OBJETIVO

Em continuidade ao trabalho desenvolvido por (NASRI, 2010) e, também, em (FERREIRA, 2013), esta tese tem como objetivo principal propor um **Método de Lagrangeano Aumentado Exponencial** para a resolução do Problema de Equilíbrio. Para tal, utilizaremos no lugar da penalidade clássica do método de lagrangeano aumentado proposto por (NASRI, 2010) , a qual só possui derivada até primeira ordem, a penalidade exponencial, a qual possui derivadas de todas as ordens.

Ainda, na tese de (FERREIRA, 2013) foram apresentadas duas abordagens de algoritmos do método de lagrangeano aumentado proposto em (NASRI, 2010). Tais algoritmos consistem na resolução de um sistema não linear e não diferenciável, originado do passo principal do algoritmo. A não a diferenciabilidade do sistema não linear é devido ao fato de que (NASRI, 2010) utlizam a penalidade quadrática clássica, conhecida na literatura, como PHR- Powell, Hestenes e Rockafellar. Por isso, a importância de se usar outras penalidades que possuam derivadas de ordem superior. É isso que estamos propondo, nesta pequisa, como uma das principais contribuições.

Com a introdução da penalidade exponencial, o sistema não linear obtido pelo **Passo 2** do algoritmo de (NASRI, 2010), será diferenciável o que passa a ser um grande ganho em relação ao trabalho apresentado em (FERREIRA, 2013). Do ponto de vista teórico,

precisamos mostrar a equivalência entre lagrangeano aumentado exponencial e método de ponto proximal.

Outras contribuições desta pesquisa são: mostrar que a sequência gerada pelo al- goritmo do novo método converge, implementar esta nova metodologia e testá-la em experimentos numéricos.

Uma consequência inevitável quando se implementam algoritmos que utilizam a penalidade exponencial é o aparecimento de problemas de mau condicionamento. Para contornar esta possível dificuldade, utilizaremos uma estratégia de interpolação quadrática na implementação do algoritmo proposto. Por fim, para realizar experimentos numéricos, escolhemos na literatura duas classes de problemas: problemas de equilíbrio geral e problemas de equilíbrio de Nash generalizados, apresentados em (FACCHINEI; KANZOW, 2009).

3 I ORGANIZAÇÃO DO TRABALHO

Neste primeiro capítulo apresentamos a motivação, objetivos e uma descrição da tese.

No segundo capítulo apresentamos um apanhado geral sobre os métodos de penalidades, com destaque para sua forma geral e estrutura algorítmica. Inicialmente, destacamos os métodos de barreiras e, em seguida, os métodos de penalidades externas. O item principal desta parte é o método de lagrangeano aumentado, aqui apresentado para problemas de otimização. Ainda neste capítulo, dada a estreita relação entre lagrangeano aumentado e método de ponto proximal para otimização em R^n, descrevemos de forma resumida este importante método, com destaque para sua estrutura algorítmica e condições de convergência. O método de ponto proximal com distância de Bregman é destacado no final.

Em seguida, no terceiro capítulo, considerando o trabalho de (NASRI, 2010), de onde retiramos o problema abordado nesta tese, nos voltamos para o problema de equilíbrio e, inicialmente, apresentamos o método de lagrangeano aumentado para este problema. Duas versões deste método são destacadas: extragaradiente inexato (IALEM) e ponto proximal extragradiente inexato (IPPEM). A versão ponto proximal extragradiente inexato irrestrito, como descrevem (NASRI, 2010), é apresentada com destaque para a fundamentação teórica que o acompanha.

No quarto capítulo, descrevemos como principal contribuição deste trabalho: um mé- todo de lagrangeano aumentado exponencial aplicado ao problema de equilíbrio, seguindo, principalmente, o exposto em (NASRI, 2010) e (FERREIRA, 2013). Na composição desta nova metodologia, destacamos dois novos Algoritmos: IALEMe e IPPEMIe. A prova de convergência é apresentada, assim como, a equivalência entre IALEMe e IPPEMIe. A estratégia de penalização exponencial poderá acarretar problemas de mau condicionamento

devido ao rápido crescimento da função exponencial. Uma alternativa para contornar esse possível problema é fazer um ajuste quadrático da função penalidade. Esse procedimento é descrito na quarta seção deste capítulo.

Com o intuito de validar a metodologia proposta, no quinto capítulo, realizamos experimentos numéricos e a análise dos resultados obtidos. Inicialmente aplicamos o Algoritmo IPPEMIe para resolver problemas de equilíbrio de Nash Generalizado e problemas de equilíbrio geral, utilizando primeiramente a penalidade exponencial e num segundo momento utilizando a penalidade exponencial com ajuste quadrático.

No último capítulo, apresentamos as considerações finais e conclusões.

LAGRANGEANO AUMENTADO PARA PROBLEMAS DE OTIMIZAÇÃO

1 I INTRODUÇÃO

Os métodos de lagrangeano aumentado são baseados em penalização e utilizados, em geral, para resolver problemas de programação não linear com restrições. São métodos que buscam resolver o problema principal através de um processo iterativo de maneira que, a cada iteração, resolve-se um problema auxiliar irrestrito. Esse subproblema é obtido adicionando-se à função objetivo um múltiplo das restrições do problema original. Dentre as referências clássicas que tratam deste método, destacamos (POWELL, 1969; HESTENES, 1969; ROCKAFELLAR, 1974; BERTSEKAS, 1976; SANTOS, 1999; IUSEM, 1995), entre outros. Uma descrição resumida de tais métodos será apresentada logo após a apresentação dos métodos de penalidades clássicos, os quais são descritos a seguir.

2 I MÉTODOS DE PENALIDADES

Nesta seção, adotaremos como referências, considerando as diversas existentes, (SAN- TOS, 1998; LUENBERGER, 2008; GRIVA; NASH; SOFER, 2009), que apresentam conceitos, algoritmos e observações sobre os métodos de penalidades. A seguir apresentamos um resumo de tais métodos.

Como tratado em (SANTOS, 1998), uma ideia bastante natural e frequentemente utilizada em programação não linear é tentar converter problemas complexos em outros, em geral mais simples, cuja resolução seja conhecida. Seria importante se um problema de otimização geral pudesse, de alguma forma, ser reduzido a um problema de minimização sem restrições, minimização em caixas ou politopos.

Na chamada penalidade interna a função objetivo é modificada agregando um termo funcional que tende a infinito quando o ponto se aproxima da fronteira do conjunto viável. Forma-se assim uma espécie de barreira intransponível: métodos irrestritos começando no interior da região são desencorajados a se aproximar do contorno devido a valores muito altos da função objetivo. Por esse motivo, os métodos de penalidade interna são também conhecidos por métodos de barreira. Esses são, por outro lado, os mais antigos métodos de pontos interiores, com prestígio radicalmente incrementado após a revolução que seguiu-se ao trabalho de (KARMARKAR, 1984).

Na penalidade externa, muitas vezes denominada simplesmente de penalidade, acrescenta-se à função objetivo um termo cujo custo aumenta com a violação das restrições. A solução de um problema penalizado externamente está, geralmente, fora do conjunto viável, mas se aproxima dele quando o termo de penalidade é muito grande.

No trabalho de (SANTOS, 1999), os autores comentam que a razão pela qual a penalidade não é o procedimento universal para lidar com restrições é que o parâmetro

que deve multiplicar a função-restrição para castigar violação (na externa) ou o risco de violação (na interna) provoca, ao tomar valores extremos, mal-condicionamento do problema penalizado. Outra dificuldade que pode surgir ao adotarmos procedimentos de penalização é que a própria estrutura do problema é transformada quando uma restrição é acrescida à função objetivo, em geral, complicando a fisionomia desta. No entanto, todas as estratégias de penalidade estão vivas na otimização atual por sua simplicidade, adaptabilidade para problemas de grande porte. Esse detalhe é fortemente destacado em (FACCHINEI; KANZOW, 2009; FACCHINEI; KANZOW, 2010b), entre outros.

Os métodos de penalidades utilizados para resolver problemas com restrições, são iterativos e cada iteração transformam o problema original em outro problema que é irrestrito. Eles servem tanto para problemas com restrição de igualdade como para restrições de desigualdade ou mesmo para aqueles que são formados por esses dois tipos de restrições. Por questões didáticas serão tratados dois tipos de métodos de penalidades: métodos de barreira e métodos de penalidade externa.

2.1 Métodos de Barreiras

Os métodos de barreira são métodos de penalidade que, em geral, são utilizados para resolver problemas com restrições de desigualdade da forma:

$$\begin{gathered} \textit{Minimizar } f(x) \\ s.a.\ g_i(x) \leq 0\ ,\ \ i = 1, 2, \ldots, m \\ x \in \mathbb{R}^n \end{gathered} \tag{2.1}$$

em que $f : \mathbb{R}^n \to \mathbb{R}$ e $g_i : \mathbb{R}^n \to \mathbb{R}$ para $i = 1, \ldots, m$, são as funções que formam o problema (2.1), sendo f a função objetivo e as funções g_i para $i = 1, \ldots, m$, as restrições que determinam o conjunto viável.

A função penalizada em relação ao problema (2.1) é definida como

$$(x, \rho) \in \mathbb{R}^n \times \mathbb{R}_{++} \mapsto B(x, \rho) = f(x) + \rho p(x) \tag{2.2}$$

em que ρ é o parâmetro de penalidade e $p : \mathbb{R}^n \to \mathbb{R}$ é uma função de penalidade, a qual é chamada, neste caso, de função barreira. Há dois tipos bastante conhecidos na literatura, a função barreira logarítmica e a função barreira inversa, as quais são definidas, respectivamente, por

$$x \in \mathbb{R}^n \mapsto p(x) = -\sum_{i=1}^{m} \ln(-g_i(x)) \tag{2.3}$$

e

$$x \in R^n \rightarrowtail p(x) = -\sum_{i=1}^{m} \frac{1}{g_i(x)} \qquad (2.4)$$

em que g_i, para $i = 1, \ldots, m$, são as restrições do problema (2.1).

A seguir apresentamos o algoritmo de penalidade interna tipo barreira.

Algoritmo 2.1 ***Algoritmo de Barreira Aplicado ao Problema*** (2.1)

Dados $\rho^0 > 0$, $k = 0$

Enquanto o critério de parada não for satisfeito, encontre:

$x^{k+1} \in \arg\min\{B(x, \rho^k) \mid x \in R^n\}$

Faça $\rho^{k+1} < \rho^k$ *(por exemplo:* $\rho^{k+1} = \frac{1}{2}\rho^k$*)*

$k = k + 1$

Continue.

No Algoritmo (2.1), B é uma função barreira (2.2), que pode ser formada pela barreira logarítmica dada na relação (2.3) ou pela barreira inversa dada na relação (2.4). A tarefa mais pesada neste algoritmo é a resolução do subproblema gerado, o qual consiste na minimização da função barreira. No entanto, este subproblema é irrestrito e, desta forma, pode ser resolvido por qualquer método de programação não linear irrestrita, como por exemplo: Newton, Gradiente Conjugado, Região de Confiança, entre outros.

Um ponto negativo dos métodos barreiras é que só se consegue provar a convergência se o parâmetro de penalidade $\rho \rightarrow 0$. Isso poderá fazer com que os subproblemas gerados pelo método tornem-se mal condicionados, como vemos no exemplo a seguir.

Considere o seguinte problema extraído de (SANTOS, 1998).

$$\begin{gathered} \textit{Minimizar } (x_1 + 1)^2 + (x_2 - 1)^2 \\ \textit{s.a. } -x_1 \leq 0 \end{gathered}$$

que tem solução $x^* = (0, 1)^T$.

Neste caso considerando a função barreira $p(x) = -\ln(x_1)$ e a função penalizada (2.2), tem-se:

$$B(x, \rho) = (x_1 + 1)^2 + (x_2 - 1)^2 - \rho \ln(x_1). \tag{2.5}$$

Para verificar o que foi dito anteriormente sobre o mau condicionamento da matriz Hessiana dos subproblemas gerados pelo Algoritmo 2.1, será calculado o gradiente e a Hessiana de B dado na relação (2.5).

$$\nabla_x B(x, \rho) = \begin{pmatrix} 2(x_1 + 1) - \rho/x_1 \\ 2(x_2 - 1) \end{pmatrix} \tag{2.6}$$

$$\nabla^2_{xx} B(x, \rho) = \begin{pmatrix} 2 + \rho/x_1^2 & 0 \\ 0 & 2 \end{pmatrix} \tag{2.7}$$

Em (2.6) os pontos estacionários, considerando $x_1 > 0$ são da forma

$$x^* = \left(\frac{-1 + \sqrt{1 + 2\rho}}{2}, \quad 1 \right).$$

E, dado que $\rho > 0$, tem-se em (2.7) que

$$\nabla^2_{xx} B(x^*, \rho) = \begin{pmatrix} 2 + \dfrac{2\rho}{\rho + 1 - \sqrt{1 + 2\rho}} & 0 \\ 0 & 2 \end{pmatrix}. \tag{2.8}$$

Como a convergência só é obtida fazendo ρ tender a zero, tem-se

$$\lim_{\rho \to 0} \frac{2\rho}{\rho + 1 - \sqrt{1 + 2\rho}} = \infty.$$

E, desta forma, a matriz $\nabla^2_{xx} B(x^*, \rho)$ em (2.8) é mal condicionada. Este foi um dos fatos que motivou o desenvolvimento de novos métodos de penalização chamados de lagrangeano aumentado.

2.2 Métodos de Penalidades Externas

Os métodos de penalizações externas são utilizados tanto para resolver problemas com restrições de desigualdade como para problemas com restrições de igualdade, ou seja, na forma

$$\begin{aligned} &\min f(x) \\ &s.a.\ g(x) \leq 0 \ e \ c(x) = 0 \end{aligned} \tag{2.9}$$

com $g : \mathbb{R}^n \to \mathbb{R}^m$ e $c : \mathbb{R}^n \to \mathbb{R}^p$.

A solução de um problema penalizado externamente está, geralmente, fora do conjunto viável, mas se aproxima dele quando o termo de penalidade é muito grande, (SANTOS, 1999).

Se considerarmos problemas contendo somente restrições de igualdade, a penalidade utilizada com mais frequência na literatura é a quadrática, apresentada a seguir:

$$x \in R^n \mapsto p(x) = \frac{1}{2}\,||c(x)||^2 = \frac{1}{2}\,c(x)^T c(x) = \frac{1}{2}\sum_{i=1}^{p}(c^i(x))^2 \qquad (2.10)$$

e a função penalidade externa neste caso será da forma:

$$(x, p) \in R^n \times R{+}{+} \mapsto P(x, p) = f(x) = \frac{\rho}{2}\sum_{i=1}^{p}(c_i(x))^2 \qquad (2.11)$$

com $\rho > 0$, fixado. Neste caso, ρ irá aumentar em todas as iterações.

Por outro lado, para problemas com restrições de desigualdade uma penalidade utilizada é a seguinte:

$$x \in R^n \mapsto p(x) = \sum_{i=1}^{m} max\{0, g_i(x)\} \qquad (2.12)$$

Podemos observar que esta penalidade é não diferenciável em pontos do seu o domínio. Portanto, uma alternativa é substituir esta penalidade por:

$$x \in R^n \mapsto p(x) = \sum_{i=1}^{m} (max\{0, g_i(x)\})^2 \qquad (2.13)$$

Neste caso, a função penalidade externa assume a forma:

$$(x, p) \in R^n \times R{+}{+} \mapsto P(x, p) = f(x) + p\sum_{i=1}^{m} (max)\{0, g_i(x)\})^2 \qquad (2.14)$$

Agora, considerando o problema (2.9), temos:

$$(x, p) \in R^n \times R{+}{+} \mapsto P(x, p) = f(x) + p\left(\sum_{i=1}^{m}(\max\{0, g_i(x)\})^2 + \frac{1}{2}\sum_{i=1}^{p}(c_i(x))^2\right) \qquad (2.15)$$

O algoritmo a seguir sintetiza o método de penalização externa aplicado ao problema (2.9).

Algoritmo 2.2 ***Penalidade Externa Aplicada ao Problema*** (2.9)

Dados: $x^0 \in Rn$, $\rho^0 > 0$

$k = 0$

Enquanto o critério de parada não for satisfeito, encontre:

$x^{k+1} \in \arg\min\{P(x, \rho^k) \mid x \in R^n\}$

Faça $\rho^{k+1} > \rho^k$ *(por exemplo:* $\rho^{k+1} = 2\rho^k$*)*

$k = k + 1$

Continue.

A penalidade $P(x, \rho^k)$ é dada em (2.15) e ρ^k é o parâmetro de penalidade.

3 | LAGRANGEANO AUMENTADO

Os métodos de lagrangeano aumentado foram inicialmente propostos, para problemas com restrições de igualdade, por (POWELL, 1969) e (HESTENES, 1969) em meados da década de 60. Foram chamados, nesta época, de métodos de multiplicadores e mais tarde, generalizados por (ROCKAFELLAR, 1974), para problemas com restrições de desigualdade. Descreveremos nesta seção os métodos de lagrangeano aumentado para problemas com restrições de igualdade e desigualdade. Tais métodos seguem a mesma filosofia dos métodos de penalidades. A diferença principal está na introdução dos multiplicadores de Lagrange, também conhecidos como variáveis duais. Desta forma, são geradas duas sequências, $\{x^k\}$, chamada de primal e, $\{\mu^k\}$, chamada de dual. A sequência primal, em geral, não é viável (não satisfaz $g(x) \leq 0$), mas espera-se que vá se tornando viável à medida que k cresce. Já a sequência dual é sempre viável, no sentido da viabilidade dual, $\mu^k \geq 0$, para todo $k \geq 0$.

Segundo (MATIOLI, 2001), uma característica importante dos métodos de lagrangeano aumentado é a relação de dualidade que estes possuem com os métodos de ponto proximal. Afirma, ainda que, um dos grandes resultados alcançados durante a época em que estes métodos foram sendo resolvidos foi exatamente esta conexão entre as sequências primal e dual ou, em termos de métodos, do lagrangeano aumentado e de ponto proximal. Ainda segundo (MATIOLI, 2001), este resultado garante que, sob hipóteses adequadas, a sequência de multiplicadores gerada pelo algoritmo de lagrangeano aumentado é a mesma obtida por um certo algoritmo de ponto proximal aplicado ao problema dual. Destaca-se, ainda, que toda a teoria de convergência da sequência de multiplicadores gerada pelos algoritmos de lagrangeano aumentado é desenvolvida no dual, via métodos proximais e a parte de computação é feita toda no primal.

Os métodos de ponto proximal fazem uma espécie de regularização somando à função objetivo um termo positivo, geralmente chamado de núcleo. Quando são aplicados ao problema dual do problema de programação convexa funcionam como uma espécie de barreira, pois forçam os pontos gerados a ficarem no ortante positivo. Veja uma descrição completa de tal método em (IUSEM, 1995).

Nos métodos de lagrangeano aumentado mais modernos, desenvolvidos nos últimos anos, praticamente toda teoria de convergência é desenvolvida utilizando métodos de ponto proximal com quase-distâncias generalizadas tipo Bregman ou ϕ-divergências. Por outro lado, a implementação destes métodos é feita no primal, isto é, os subproblemas resolvidos envolvem somente variáveis primais estando as variáveis duais fixas, como será visto adiante.

Trataremos a seguir, em seções distintas, os métodos de lagrangeano aumentado

para problemas com restrições de igualdade e restrições de desigualdade.

3.1 Lagrangeano Aumentado para Problemas com Restrições de Igualdade

Nesta subseção apresentamos o método de lagrangeano aumentado em sua forma clássica, ou seja, em sua formulação semelhante àquela dada inicialmente por (POWELL, 1969) e (HESTENES, 1969).

O problema a ser tratado é da forma:

$$\begin{aligned} &Minimizar\ f(x) \\ &s.a.\ c_i(x) = 0\ ,\ \ i = 1, \ldots, p \\ &x \in \mathbb{R}^n \end{aligned} \tag{2.16}$$

em que $f : \mathbb{R}^n \to \mathbb{R}$ e $c_i : \mathbb{R}^n \to \mathbb{R}$ para $i = 1, \ldots, p$, são funções continuamente diferenciáveis.

A função lagrangeana associada ao problema (2.16) é definida por:

$$(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^p \mapsto l(x, \lambda)\ f(x) + \lambda^T c(x) = f(x) + \sum_{i=1}^{p} \lambda_i C_i(x) \tag{2.17}$$

As condições necessárias de primeira ordem garantem que se x^* é uma solução ótima do problema (2.16), então existe um vetor de multiplicadores de Lagrange λ^* tal que $\nabla_x l(x^*, \lambda^*) = 0$, em que l é a função lagrangeana (2.17).

Desta forma, se f e c_i, para todo $i = 1, \ldots, p$, são diferenciáveis então os candidatos a pontos estacionários para o problema (2.16) devem satisfazer o seguinte sistema de $n + p$ equações não lineares por $n + p$ variáveis

$$\begin{cases} \nabla f(x) + \lambda^T \nabla c(x) = 0 \\ c_i(x) = 0 \quad i = 1, 2, \ldots, p. \end{cases} \tag{2.18}$$

Veremos que os multiplicadores de Lagrange no método de lagrangeano aumentado a ser apresentado nesta seção são atualizados de maneira que a primeira equação em (2.18) seja satisfeita.

Definição 2.1 *Dizemos que dois problemas A e B são equivalentes quando toda solução de A* é *também uma solução de B e vice-versa.*

No sentido da Definição 2.1 podemos verificar que o problema (2.16) é equivalente ao seguinte problema

$$Minimizar\ f(x) + \lambda^T c(x)$$

$$s.a.\ c_i(x) = 0\ ,\ \ i = 1, \ldots, p \qquad (2.19)$$
$$x \in \mathbb{R}^n$$

Utilizando a penalidade quadrática apresentada em (2.10), tem-se a seguinte função penalizada para o problema (2.19)

$$(x, \lambda, \rho) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}_{++} \mapsto L(x, \lambda, \rho) = f(x) + \lambda^T c(x) + \frac{\rho}{2} c^T(x) c(x) =$$
$$= f(x) + \sum_{i=1}^{p} \lambda_i c_i(x) + \frac{\rho}{2} \sum_{i=1}^{p} (c_i(x))^2. \qquad (2.20)$$

Neste caso, ρ é o parâmetro de penalidade e λ é o vetor de multiplicadores de Lagrange associado às restrições de igualdade. Pode-se observar que na função (2.20) aparecem p variáveis a mais que na função penalidade externa (2.11) apresentada anteriormente neste capítulo. No entanto, com a abordagem proposta por (POWELL, 1969) e (HESTENES, 1969) estas variáveis são atualizadas por uma fórmula fechada. Assim, para determinar uma boa estimativa para os multiplicadores de Lagrange, deriva-se a função L dada em (2.20) com relação à variável x , estando λ e ρ fixados, e iguala-se a zero, ou seja,

$$\nabla L(x, \lambda, \rho) = \nabla f(x) + \lambda^T \nabla c(x) + \rho c^T(x) \nabla c(x) =$$
$$= \nabla f(x) + (\lambda^T + \rho c(x))^T \nabla c(x) = 0. \qquad (2.21)$$

Desta forma, uma boa estimativa para os multiplicadores é

$$\bar{\lambda} = \lambda + \rho c(x), \qquad (2.22)$$

pois (2.21) teria a forma

$$\nabla f(x) + \bar{\lambda} \nabla^T c(x) = 0 \qquad (2.23)$$

e esta última equação é exatamente a primeira equação dada na relação (2.18), ficando apenas a segunda equação, $c(x) = 0$ ($c(x) = (c_1(x), \ldots, c_p(x))^T$, para ser satisfeita.

Logo, a metodologia proposta por Powell e Hestenes é a introdução de um algoritmo que gere candidatos à solução, no sentido de serem pontos estacionários do problema (2.16). Estes devem satisfazer o sistema (2.18). Portanto, atualizando-se os multiplicadores de Lagrange pela fórmula (2.22) a primeira equação do sistema (2.18) é satisfeita automaticamente, enquanto a segunda é forçada a ser satisfeita através da atualização do parâmetro de penalidade.

Posto isso, apresentamos o algoritmo de lagrangeano aumentado proposto por (POWELL, 1969) e (HESTENES, 1969).

Algoritmo 2.3 Lagrangeano Aumentado Aplicado ao Problema com Restri- ção de Igualdade

Dados $\rho^0 > 0$, $\lambda^0 \in R^m$, $\alpha > 1$, $k = 0$

Enquanto o critério de parada não for satisfeito, encontre:

$x^{k+1} \in \arg\min\{L(x, \lambda^k, \rho^k) \mid x \in R^n\}$, *com L dada em* (2.20) (2.24)

Faça: $\lambda^{k+1} = \lambda^k + \rho^k c(x^{k+1})$

Se $\|c(x^{k+1})\| > 0,1\|c(x^k)\|$

Faça $\rho^{k+1} = \alpha\rho^k$ *(por exemplo: $\alpha = 2$).*

Senão $\rho^{k+1} = \rho^k$

$k = k + 1$

Continue.

Alguns comentários sobre o algoritmo: o subproblema (2.24), ou seja, dado $\lambda^k \in R^p$ e $\rho^k > 0$, encontrar

$$x^{k+1} \in \arg\min\{L(x, \lambda^k, \rho^k) : x \in R^n\}$$

é chamado subproblema interno, que consiste na minimização de uma função irrestrita e a função L é o lagrangeano aumentado.

A parte representativa de maior esforço computacional do algoritmo é encontrar uma solução para este subproblema, o que pode ser executado utilizando-se métodos de programação irrestrita, como Newton, Região de Confiança, dentre outros.

Após a resolução do subproblema (2.24) atualizam-se os multiplicadores de Lagrange λ, pela relação (2.22). Em seguida faz-se uma verificação da viabilidade, $c(x) = 0$. Se houver ganho de viabilidade do ponto x^k para o ponto x^{k+1} mantém-se o parâmetro de penalidade, caso contrário, o mesmo é aumentado. Continua-se o processo até que algum critério de parada seja satisfeito.

Como o parâmetro de penalidade só é atualizado se as restrições estiverem "pouco satisfeitas", o que não ocorre nos métodos de penalidades clássicos apresentados nas

seções precedentes, reduz-se a possibilidade de mau condicionamento dos subproblemas gerados pelo método, o que os torna vantajosos.

3.2 Lagrangeano Aumentado para Problemas com Restrições de Desigualdade

Na sequência será apresentado o método de lagrangeano aumentado proposto em (ROCKAFELLAR, 1973) (ver também (MINOUX, 1986) e (BERTSEKAS, 1976)), para o problema com restrições de desigualdade, generalizando desta forma o método de multiplicadores proposto por (POWELL, 1969) e (HESTENES, 1969), apresentado na subseção anterior. Além disso, serão apresentados métodos de lagrangeano aumentado mais modernos cuja teoria de convergência é feita utilizando-se o método de ponto proximal aplicado ao problema dual com alguma quase-distância (Bregman ou ϕ-divergência), seguindo o exposto em, por exemplo, (IUSEM, 1995; IUSEM; TEBOULLE, 1995).

O problema tratado nesta subseção é com restrições de desigualdade e tem a forma

$$\begin{array}{c} \text{Minimizar } f(x) \\ \text{s.a. } g_i(x) \leq 0\ ,\ \ i = 1, 2, \ldots, m, \\ x \in \mathbb{R}^n \end{array} \qquad (2.25)$$

em que, f e g_i, para $i = 1, 2, \ldots, m$, são funções definidas em $\mathbb{R}^n$ com valores em R.

Uma opção para a generalização do método de lagrangeano aumentado para o problema (2.25) é obtida através da introdução de variáveis de folga para as restrições de desigualdade $g(x) \leq 0$, onde $(g(x) = (g_1(x), g_2(x), \ldots, g_p(x))^T$, conforme encontramos em (MINOUX, 1986; ROCKAFELLAR, 1973). Desta forma, o problema (2.25), com variáveis de folga $s_i \geq 0$, $i = 1, \ldots, m$, possui o formato:

$$\begin{array}{c} \text{Minimizar } f(x) \\ \text{s.a. } g_i(x) + s_i = 0\ ,\ \ i = 1, 2, \ldots, m. \\ s_i \geq 0\ ,\ i = 1, 2, \ldots, m \\ x \in \mathbb{R}^n \end{array} \qquad (2.26)$$

Como foi feito na seção anterior, ao elaborar um algoritmo de lagrangeano aumentado aplicado ao problema (2.26), o subproblema gerado tem a forma:

$$\text{Minimizar } \left\{ f(x) + \sum_{i=1}^{m} \mu_i (g_i(x) + s_i) + \frac{\eta}{2} \sum_{i=1}^{m} (g_i(x) + s_i)^2 \ :\ x \in \mathbb{R}^n\ ,\ s \geq 0 \right\}, \quad (2.27)$$

e, neste caso, μ é o vetor de multiplicadores de Lagrange e η é o parâmetro de penalidade.

Para o parâmetro de penalidade positivo, mostra se que o problema (2.27) se reduz

à (ver (ROCKAFELLAR, 1974; MINOUX, 1986)):

$$\text{Minimizar} \quad \left\{ f(x) + \sum_{i=1}^{m} P(g_i(x), \mu_i, \eta) \ : \ x \in \mathbb{R}^n \right\}, \tag{2.28}$$

e, neste caso, para $i = 1, \ldots, m$, a função penalidade P é dada pela fórmula:

$$P(g_i(x), \mu_i, \eta) = \begin{cases} \eta(g_i(x))^2 + \mu_i g_i(x) & se \ g_i(x) \geq \dfrac{-\mu_i}{2\eta}. \\ \dfrac{-\mu_i^2}{4\eta} & c.c. \end{cases} \tag{2.29}$$

Note que a variável de folga s não aparece em (2.28) e nem em (2.29). Portanto, o algoritmo de lagrangeano aumentado desenvolvido para o problema com restrições de desigualdade é análogo ao Algoritmo 2.3, porém com subproblema interno dado por (2.28) e (2.29). Desta forma, substituindo a penalidade P dada em (2.29) na função (2.28) temos a função lagrangeano aumentada proposta por Rockafellar

$$(x, \mu, \eta) \in \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}_{++} \mapsto L(x, \mu, \eta) = f(x) + \frac{1}{4\eta} \sum_{i=1}^{m} \left[(\max\{0, \mu_i + 2\eta g_i(x)\})^2 - \mu_i^2 \right] \tag{2.30}$$

Algoritmo 2.4 Lagrangeano Aumentado para o Problema com Restrição de Desigualdade

Dados $\mu^0 \in \mathbb{R}_+^m$, $\{\eta_k\}$ *uma sequência de termos positivos e* $k = 0$.

Enquanto o critério de parada não for satisfeito, encontre:

$$x^{k+1} = \arg\min\{L(x, \mu^k, \eta_k) \mid x \in \mathbb{R}^n\} \ \textit{com L dada em } (2.30)) \tag{2.31}$$

Faça:

$$\mu_i^{k+1} = \max\{0, \mu_i^k + 2\eta_k g_i(x^{k+1})\}, \qquad para \ i = 1, \ldots, m$$

$k = k + 1$

Continue.

3.3 O MÉTODO DE PONTO PROXIMAL PARA OTIMIZAÇÃO EM $\mathbb{R}^n$

Na seção anterior descrevemos um método de lagrangeano aumentado aplicado ao problema de otimização com restrições de desigualdade, utilizando variáveis de folga para transformar as restrições do problema (2.25) em igualdades. Na sequência, descrevemos outra metodologia baseada nos métodos de ponto proximal, conforme descrito em (IUSEM,

1995).

A ideia é baseada nos métodos primais-duais, ou seja, apresentaremos a convergência do método de ponto proximal para um problema irrestrito e depois, a equivalência entre as sequências duais geradas pelos métodos de ponto proximal e lagrangeano aumentado.

Primeiramente, descrevemos o método de ponto proximal aplicado ao seguinte problema irrestrito

$$\begin{aligned} &Minimizar\ f(x) \\ &s.a.: x \in \mathbb{R}^n \end{aligned} \tag{2.32}$$

no qual $f : \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função convexa e limitada por baixo.

Mesmo o problema sendo convexo e f limitada por baixo, não existe a garantia de um único minimizador. Para assegurar a unicidade do minimizador é acrescentado à função objetivo do problema (2.32) um termo de regularização. Assim, o método de ponto proximal gera uma sequência $\{x^k\}$ da seguinte maneira:

$$x^0 \in \mathbb{R}^n \tag{2.33}$$

$$x^{k+1} = argmin_{x \in \mathbb{R}^n} \{f(x) + \alpha_k \|x - x^k\|^2\} \tag{2.34}$$

em que α_k é um número real satisfazendo $0 < \alpha_k < \bar{\alpha}$ para algum $\bar{\alpha} > 0$, incluindo o caso α_k constante.

O resultado a seguir estabelece a convergência da sequência $\{x^k\}$, gerada pelo método de ponto proximal (2.33) e (2.34).

Teorema 2.1 *Convergência do método de ponto proximal*

Considere $f : \mathbb{R}^n \rightarrow \mathbb{R}$ *uma função convexa e continuamente diferenciável. Assuma que o conjunto solução U de minimizadores de f em* $\mathbb{R}^n$ é não vazio. Então a sequência gerada *por* (2.33) *e* (2.34) *converge para um ponto* $\bar{x} \in U$.

Prova. Ver (IUSEM, 1995).

Para provar a convergência da sequência dual $\{\mu^k\}$, gerada pelo Algoritmo 2.4, será necessário introduzir o problema dual associado ao problema primal (2.25). Para isso, considere f e g_i convexas e diferenciáveis, para todo $i = 1, 2, \ldots, m$. Associado ao problema (2.25) define-se a função lagrangeana clássica $l : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ por

$$\ell(x,\mu) = \begin{cases} f(x) + \sum_{i=1}^{m} \mu_i g_i(x) & \text{se } \mu \geq 0 \\ +\infty & \text{c.c.} \end{cases} \tag{2.35}$$

Considere a função objetivo $\psi : R^m \in R$ definida por

$$\psi(\mu) = \min_{x \in R^n} l(x, \mu) \qquad (2.36)$$

em que l é a função lagrangeana (2.35).

O resultado seguinte prova que a sequência $\{\mu^k\}$, gerada pelo Algoritmo 2.4, é a mesma sequência gerada pelo algoritmo de ponto proximal (2.33) e (2.34) quando aplicado ao problema $-\psi$.

Teorema 2.2 *Equivalência entre ponto proximal e lagrangeano aumentado - para problema de otimização*

Considere $\{\bar{\mu}^k\}$ *a sequência gerada pelo algoritmo de ponto proximal,* (2.33) *e* (2.34)*, aplicado ao problema* $\min_{\mu \in R^m} -\psi(\mu)$*, com* $\alpha_k = \frac{1}{4\eta_k}$ *e* $\{\mu^k\}$ *a sequência dual gerada pelo algoritmo de lagrangeano aumentado 2.4. Se* $\bar{\mu}^0 = \mu^0$ *então* $\bar{\mu}^k = \mu^k$ *para todo k.*

Prova. Ver (IUSEM, 1995).

Até este momento poderíamos nos dar por satisfeitos, pois como mencionado nos parágrafos precedentes, os métodos de lagrangeano aumentado para os problemas com restrições de igualdade e desigualdade apresentados até aqui, contornam, de certa forma, o problema do mau condicionamento gerado pelos métodos de penalização clássicos.

Com a introdução dos multiplicadores, ou variáveis duais, foi possível mostrar a convergência (dual no caso de desigualdade) dos métodos de lagrangeano aumentado, sem que o parâmetro de penalidade aumentasse em todas as iterações. Inclusive, este parâmetro pode ser constante para o método de lagrangeano aumentado, proposto por (ROCKAFELLAR, 1974), para problemas com restrição de desigualdade como pode ser visto no Algoritmo 2.4.

A questão principal está na determinação da solução do subproblema gerado pelo Algoritmo 2.4, ou seja, na solução do subproblema dado na relação (2.30). Note que esta função só possui derivada, em todo o seu domínio de definição, até primeira ordem. Portanto, métodos que usam informação de segunda ordem, como é o caso do método de Newton e de região de confiança, não podem ser utilizados, ou se forem utilizados podem gerar instabilidade quando os pontos obtidos se aproximarem de valores em que a função não é diferenciável. Desta forma, a comunidade de otimização desenvolveu métodos em que as funções envolvidas possuam derivadas de ordem superiores.

O desenvolvimento seguinte visa explorar a relação primal-dual. Mostraremos a convergência do método de ponto proximal e, em seguida, a equivalência com o método de lagrangeano aumentado. Porém agora, ao invés da regularização quadrática apresentada

na função (2.34) do método de ponto proximal, utilizaremos uma quase distância mais geral. Da mesma forma, no primal deveremos substituir a penalidade quadrática da função lagrangeano aumentado (2.30) por uma quase distância que possua derivada de ordens superiores.

As quase-distâncias mais utilizadas são: distâncias de Bregman e ϕ-divergência. Nesta tese, em função de sua especificidade, será apresentado somente o caso de distâncias de Bregman. Para isso, o desenvolvimento apresentado está baseado em (IUSEM, 1995).

Diferentemente da regularização utilizada no método de ponto proximal quadrático, aqui será necessário que este termo penalize pontos inviáveis, pois o interesse é aplicar o método de ponto proximal ao problema dual do problema (2.25), no qual as variáveis duais são não negativas. Sendo assim, a regularização funcionará como nos métodos de barreira, descritos no início deste capítulo. O objetivo da regularização é penalizar "infinitamente" pontos que se aproximem da fronteira. Para isso, considere o seguinte problema restrito:

$$\begin{aligned} &\textit{Minimizar } f(x) \\ &\textit{s.a.} \quad x \in \bar{S} \end{aligned} \tag{2.37}$$

em que $\bar{S} \subset \mathrm{R}^n$ é fechado.

Seguindo o exposto em (IUSEM, 1995), uma abordagem para resolver (2.37) é a introdução de uma função de penalidade p satisfazendo $p(x) = \infty$ se $x \notin S$. Para $\alpha > 0$, o problema

$$\min_{x \in \mathrm{R}^n} \{f(x) + \alpha p(x)\} \tag{2.38}$$

terá sua solução em $\bar{S}$ (se as soluções existirem), como foi o caso desenvolvido para o método de ponto proximal clássico. Se (2.37) tiver solução e p for apropriadamente escolhida, então (2.38) terá uma única solução $x(\alpha)$ e sob hipóteses adicionais pode ser provado que $\lim_{\alpha \to 0} x(\alpha)$ existe e resolve (2.37).

Como no caso da regularização (aqui olhando também como penalização clássica), o problema é que (2.38) se torna mal condicionado quando α se torna muito pequeno, isto é, para pontos próximos da fronteira de $\bar{S}$, $p(x)$ se torna muito grande e $\alpha p(x)$ se torna indeterminado quando α se aproxima de 0. Logo, a ideia é fazer uma combinação de penalização com a abordagem de ponto proximal para obter convergência para α distante de zero. Para alcançar este objetivo, usaremos uma função quase-distância $D : S \times S \to \mathrm{R}+$ (em que S é o interior de $\bar{S}$) tal que $D(x, y) = 0$ se $x = y$ e se aproxima de infinito quando y se aproxima da fronteira ∂S de $\bar{S}$.

3.4 Funções e Distâncias de Bregman

Considere $S \subset R^n$ não vazio, aberto, convexo e seu fecho $\bar{S}$, d uma função convexa real em $\bar{S}$ e $D_d : \bar{S} \times S \to R$ definida por

$$D_d(x, y) = d(x) - d(y) - \nabla d(y)^T (x - y). \tag{2.39}$$

A função d será chamada uma função de Bregman e D_d uma distância de Bregman induzida por d, se cumprir as seguintes condições:

B1) d é continuamente diferenciável em S.

B2) d é estritamente convexa e contínua em $\bar{S}$.

B3) Para $\delta \in R$, os conjuntos de nível parciais

$$\Gamma_1(x, \delta) = \{x \in \bar{S} : D_d(x, y) \leq \delta\} \text{ e } \Gamma_2(x, \delta) = \{y \in S : D_d(x, y) \leq \delta\},$$

são limitados para todo $y \in S$ e todo $x \in \bar{S}$, respectivamente.

B4) Se $\{y^k\} \subset S$ convegir para y^* então $D_d(y^*, y^k)$ converge para 0.

B5) Se $\{x^k\} \subset \bar{S}$ e $\{y^k\} \subset S$ são sequências tais que $\{x^k\}$ é limitada, $\lim_{k \to \infty} y^k = y^*$ e $\lim_{k \to \infty} D_d(x^k, y^k) = 0$ então $\lim_{k \to \infty} x^k = y^*$.

De acordo com (IUSEM, 1995), vale a pena destacar algumas observações a respeito da definição e propriedades anteriores:

(i) O conjunto S é chamado de zona de d.

(ii) $D_d(x, y) \geq 0$ para todo $x \in \bar{S}$ e $y \in S$. Além disso, $D_d(x, y) = 0$ se, e somente se, $x = y$.

(iii) As propriedades **B**4 e **B**5 precisam ser verificadas somente nos pontos da fronteira ∂S de S, pois quando x^k e y^* estão em S elas se verificam automaticamente, como uma consequência de **B**1, **B**2 e **B**3.

Uma função de Bregman d será chamada de coerciva na fronteira se a seguinte propriedade for satisfeita:

B6) Se $\{y^k\} \subset S$ é tal que $\lim_{k \to \infty} y^k = y \in \partial S$, então $\lim_{k \to \infty} \nabla d(y^k)^T (x - y^k) = -\infty$ para todo $x \in S$.

A seguir são apresentados alguns exemplos de funções de Bregman que podem ser encontrados em (IUSEM, 1995).

Exemplo 2.1 $S = R^n$, $d(x) = x^t M x$ *com* $M \in R^{n \times n}$ *simétrica e positiva definida. Neste caso,* $D_d(x, y) = (x - y)^T M (x - y) = \|x - y\|^2_M$ ($\|.\|_M$ *é a norma induzida por M).*

Exemplo 2.2 $S = \mathbb{R}^n_{++}$, $d(x) = \sum_{j=1}^{n} x_j \log x_j$ *extendida com continuidade em* $\partial\mathbb{R}^n_+$ *e convenção que* 0 log 0 = 0. *Neste caso,* $D_d(x, y) = \sum_{j=1}^{n}(x_j \log \frac{x_j}{y_j} + y_j - x_j)$, *a qual é conhecida na literatura como entropia (ou divergência) de Kullback-Leibler.*

Exemplo 2.3 $S = \mathbb{R}^n_{++}$, $d(x) = \sum_{j=1}^{n}(x_j^\alpha - x_j^\beta)$ *com* $\alpha \geq 1$, $0 < \beta < 1$. *Para* $\alpha = 2$ *e* $\beta = 1/2$ *tem se* $D_d(x, y) = ||x - y|| + \sum_{j=1}^{n} \frac{1}{2\sqrt{y_j}}(\sqrt{x_j} - \sqrt{y_j})^2$ *e para* $\alpha = 1$ *e* $\beta = 1/2$ $D_d(x, y) = \sum_{j=1}^{n} \frac{1}{2\sqrt{y_j}}(\sqrt{x_j} - \sqrt{y_j})^2$.

3.5 O Método de Ponto Proximal com Distância de Bregman

Considere $S \subset \mathbb{R}^n$ aberto e convexo, $\bar{S}$ o fecho de S e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função convexa e contínua em $\bar{S}$. O método de ponto proximal aplicado ao problema (2.37) é definido por:

$$x^0 \in S \tag{2.40}$$

$$x^{k+1} = argmin_{x \in \bar{S}}\{f(x) + \alpha_k D_d(x, x^k)\} \tag{2.41}$$

em que d é uma função de Bregman com zona S e $0 < \alpha_k < \bar{\alpha}$ para algum $\bar{\alpha} > 0$.

A seguir trataremos da convergência do método de ponto proximal. Para isso assumimos que a função f que define o problema (2.37) é limitada por baixo em $\bar{S}$.

Teorema 2.3 *Convergência do método de ponto proximal com distância de Bregman. Se o problema* (2.37) *tem soluções e* d é *coerciva na fronteira com respeito a S, então a sequência* $\{x^k\}$ *gerada por* (2.40) *e* (2.41) *converge para uma solução* x^* *do problema* (2.37).

Prova. Ver (IUSEM, 1995).

Apesar da existência de uma gama de metodologias de lagraneano aumentado utilizando distâncias de Bregman, aqui trataremos somente o caso do método exponencial apresentado por Bertsekas conforme (IUSEM, 1995) (ver também (BERTSEKAS, 1976)), o qual será de grande valia para nos auxiliar no desenvolvimento da nossa proposta, a ser apresentada no próximo capítulo.

Considere novamente o problema de otimização com restrições de desigualdade, ou seja

$$\begin{aligned} &Minimizar\ f(x) \\ &s.a.\ g_i(x) \leq 0\ ,\ \ i = 1, 2, \ldots, m, \\ &x \in \mathbb{R}^n \end{aligned} \tag{2.42}$$

no qual f e g_i, para $i = 1, 2 \ldots, m$, são funções convexas e diferenciáveis.

A função lagrangeano aumentado com penalidade exponencial, associada ao problema (2.42), é dada por:

$$(x, \mu, \eta) \in R^n \times R^m_+ \times R_{++} \rightarrowtail L(x, \mu, \eta) = f(x) + \frac{1}{\eta}\sum_{i=1}^{m} \mu_i e^{\eta g_i(x)} \qquad (2.43)$$

onde $\eta > 0$ é o parâmetro de penalidade e μ é o vetor dos multiplicadores de Lagrange.

Note que a função lagrangeano aumentado exponencial (2.43) é tantas vezes diferenciável quanto forem f e g_i. O mesmo não se verifica no caso da função lagrangeano aumentado quadrática de Rockafellar considerada na relação (2.30).

Fixada uma sequência de termos não negativos $\{\eta_k\}$, o método de lagrangeano aumentado exponencial gera as sequências $\{x^k\}$ e $\{\mu^k\}$, através dos seguintes passos:

$$\mu^0 \in R^m_+ \text{ e para } k = 0, 1, 2, \ldots \qquad (2.44)$$

$$x^{k+1} = argmin_{x \in R^n} \{L(x, \mu^k, \eta_k)\} \qquad (2.45)$$

$$\mu_i^{k+1} = \mu_i^k e^{\eta_k g_i}(x^{k+1}), \quad i = 1, 2, \ldots, m \qquad (2.46)$$

em que L é a função lagrangeano aumentado exponencial (2.43).

Equivalentemente ao Teorema 2.2, é possível mostrar a convergência da sequência dual, $\{\mu^k\}$, gerada pelo método de lagrangeano aumentado exponencial, através da equivalência entre ponto proximal com distância de Bregman e lagrangeano aumentado. Para isso, recorremos ao seguinte resultado:

Teorema 2.4 *Teorema de equivalência para o problema de otimização com Bregman Considere $\{\mu^k\}$ a sequência gerada por* (2.44)-(2.46) *e $\{\overline{\mu}k\}$ a sequência obtida através de*

$$\bar{\mu}^{k+1} = argmin_{\mu \geq 0}\{-\psi(\mu) + \frac{1}{\eta_k} D_d(\mu, \mu^k)\} \qquad (2.47)$$

com $\psi(\mu) = \min_{x \in \mathbb{R}^n}\{f(x) + \sum_{i=1}^{m} \mu_i g_i(x)\}$ e d: $R^m_+ \rightarrow R$ dada por d(x) = $\sum_{j=1}^{m} x_j \log x_j$. *Se* $\mu^0 = \overline{\mu}^0$ *então* $\mu^k = \overline{\mu}^k$ *para todo k.*

Prova. Ver (IUSEM, 1995).

Note que, na equação (2.47), foi necessário escolher o parâmetro de regularização

$\alpha_k = \frac{1}{\eta^k}$, sendo η_k o parâmetro de penalidade no método de lagrangeano aumentado exponencial. Esta escolha é necessária para se conseguir demonstrar tal equivalência.

LAGRANGEANO AUMENTADO PARA PROBLEMAS DE EQUILÍBRIO

1 I INTRODUÇÃO

Atualmente o problema de equilíbrio tem recebido muita atenção dos pesquisadores, comprovado pelos vários resultados teóricos e experimentos publicados. Entre estes estudos destacamos: (BIANCHI; PINI, 2005; BIGI et al., 2013; BLUM; OETTLI, 1994; BURACHIK, 2012; FACCHINEI, 2003; FLåM; ANTIPIN, 1996; IUSEM; NASRI, 2007; NASRI, 2010; IUSEM; KASSAY; SOSA, 2009; IUSEM; SOSA, 2003a; IUSEM; SOSA, 2003b; MASTROENI, 2000; MATIOLI; SOSA; YUAN, 2012; MUU; OETTLI, 1992; MUU; QUOC, 2009; FACCHINEI; KANZOW, 2007; QUOC; MUU, 2004; TRAN; DUNG; NGUYEN, 2008; SANTOS; SCHEIMBERG, 2011; NGUYEN; STRODIOT; NGUYEN, 2009; NGUYEN; STRODIOT et al., 2009). Em particular (BIGI et al., 2013), faz uma descrição da situação atual no que se refere a algoritmos e condições de existência de soluções.

O problema de equilíbrio fornece uma abordagem unificadora de vários problemas matemáticos. Entre estes, inclui como casos particulares, problemas de otimização convexa, problemas de ponto fixo, problemas de complementaridade, problemas de equilíbrio de Nash, problemas de desigualdade variacional e problemas de minimização vetorial, como destacado em (BIGI et al., 2013; BLUM; OETTLI, 1994; IUSEM; SOSA, 2003b).

Neste capítulo apresentaremos o problema de equilíbrio e uma abordagem recente apresentada por (NASRI, 2010), sobre o método de lagrangeano aumentado para a sua resolução. Neste artigo, os autores apresentam métodos de lagrangeano aumentado para resolução de problemas de equilíbrio em dimensão finita, cujos conjuntos viáveis são definidos por inequações convexas, generalizando o método de lagrangeano aumentado para otimização restrita. A cada iteração as variáveis primais são atualizadas resolvendo-se um problema de equilíbrio irrestrito, enquanto as variáveis duais são atualizadas por uma fórmula fechada. Além disso, é realizada a análise de convergência do algoritmo.

O problema de equilíbrio com as propriedades mínimas necessárias para o desenvolvimento teórico é apresentado na definição a seguir.

Definição 3.1 *Considere* K *um subconjunto de* $\mathbb{R}^n$ *não vazio, convexo e fechado e uma função* $f : K \times K \to \mathbb{R}$ *satisfazendo*

P1: $f(x, x) = 0$ *para todo* $x \in K$,

P2: $f(x, \cdot) : K \to \mathbb{R}$ é *convexa e semicontínua inferior para todo* $x \in K$,

P3: $f(\cdot, y) : K \to \mathbb{R}$ é *semicontínua superior para todo* $y \in K$.

O problema de equilíbrio, denotado por $EP(f, K)$, *consiste em determinar* $x^* \in K$ *tal que* $f(x^*, y) \geq 0$ *para todo* $y \in K$.

O conjunto de soluções para *EP* (f, K) será denotado por $S(f, K)$.

Em termos de métodos computacionais para problemas de equilíbrio, (BIGI et al., 2013; FACCHINEI; KANZOW, 2010a) comentam que existem poucas referências encontradas na literatura. Este foi um dos fatos que serviu de motivação para buscarmos o desenvolvimento de algoritmos para resolução destes problemas.

Neste cenário, destacamos o artigo de (NASRI, 2010), no qual propõem versões exatas e inexatas de métodos de lagrangeano aumentado para resolução do *EP* (f, K) em R^n, para o caso nos quais o conjunto viável K é da forma:

$$K = \{x \in R^n : h_i(x) \leq 0 \qquad (1 \leq i \leq m)\} \tag{3.1}$$

em que $h_i : R^n \rightarrow R$ são funções convexas para todo $i = 1, \ldots, m$.

Observe, na definição do problema de equilíbrio, a dependência das duas variáveis, x e y, estarem no conjunto convexo K, ou seja, quer se encontrar x^* que **e**steja em K tal que $f(x^*, y) \geq 0$, **p**ara todo y em K. Devido ao fato das variáveis estarem no conjunto convexo K, a maioria dos artigos encontrados na literatura acabam fazendo algum tipo de projeção. É bem conhecido que fazer projeção reduz-se em outro problema de otimização que pode ser de difícil solução, a menos que o conjunto K seja tipo poliedros (descritos por restrições lineares ou limites das variáveis). Com a introdução do método de lagrangeano aumentado esta dependência pode ser enfraquecida através da abordagem de penalização.

2 | LAGRANGEANO AUMENTADO APLICADO AO PROBLEMA DE EQUILÍBRIO

Assumiremos, como em (NASRI, 2010), que o conjunto K que define o problema de equilíbrio, satisfaz a Condição de Qualificação de Slater dada na definição a seguir.

Definição 3.2 *Condição de Qualificação de Slater*

Dizemos que o conjunto K, dado em (3.1)*, satisfaz a Condição de Qualificação de Slater se existir* $x \in R^n$*, tal que* $h_i(x) \leq 0$ *para* $\mathbf{i} \in \mathbf{I}$ *e* $h_i(x) < 0$ *para* $i \notin \mathbf{I}$*, sendo* $\mathbf{I}$ *o conjunto de* índices *(possivelmente vazio), tal que a função* h_i é *afim.*

Vale ressaltar que existem outras condições de qualificação que poderiam ser utilizadas.

Para maiores detalhes consultar, por exemplo, (EUSTÁQUIO, 2007).

A bifunção lagrangeana $L : (R^n \times R^m) \times (R^n \times R^m) \rightarrow R$, associada ao problema de equilíbrio *EP* $(f ; K)$, (3.1), é definida por:

$$\mathcal{L}((x,\lambda),(y,\mu)) = f(x,y) + \sum_{i=1}^{m} \lambda_i h_i(y) - \sum_{i=1}^{m} \mu_i h_i(x). \tag{3.2}$$

Destacamos que o problema de otimização (2.42) pode ser tratado como um caso particular do problema de equilíbrio *EP* (*f*, *K*), tomando-se $f(x, y) = f(y) - f(x)$. Nesse caso, (3.2) reduz-se à $L((x, \lambda), (y, \mu)) = l(y, \lambda) - l(x, \mu)$, sendo l a função lagrangeana (2.35) associada ao problema (2.42).

Na sequência, apresentaremos o método de lagrangeano aumentado proximal para *EP* (*f*, *K*), conforme apresentado em (NASRI, 2010). Considere K o conjunto dado na definição (3.1), a função penalidade $p_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m_+ \times \mathbb{R}_{++} \to \mathbb{R}$ definida por:

$$p_i(x, y, \lambda, \gamma) = \frac{\gamma}{2}\left[\left(\max\left\{0, \lambda_i + \frac{h_i(y)}{\gamma}\right\}\right)^2 - \left(\max\left\{0, \lambda_i + \frac{h_i(x)}{\gamma}\right\}\right)^2\right] \qquad (3.3)$$

e a função lagrangeano aumentado proximal $L : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}_{++} \to \mathbb{R}$ definida por:

$$L(x, y, \lambda, z, \gamma) = f(x, y) + \gamma\langle x - z, y - x\rangle + \sum_{i=1}^{m} p_i(x, y, \lambda, \gamma). \qquad (3.4)$$

Note que a função lagrangeano aumentado (3.4) torna o problema de equilíbrio *EP* (*f, K*) irrestrito, ou seja, as restrições que descrevem o conjunto *K* são incorporadas nesta. Desta forma, determina-se uma solução para o problema de equilíbrio restrito através da resolução de uma sequência de problemas irrestritos, utilizando-se a metodologia delagrangeano aumentado proximal.

A seguir descrevemos os dois principais algoritmos apresentados em (NASRI, 2010) para o problema de equilíbrio *EP* (*f, K*): lagrangeano aumentado ex- tragradiente inexato e ponto proximal extragradiente inexato.

Algoritmo 3.1 ***Lagrangeano Aumentado Extragradiente Inexato (IALEM)***

Passo 1: *Considere uma sequência limitada* $\{\gamma_k\} \subset \mathbb{R}_{++}$, $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m_+$, *uma tolerância* $\delta \in (0, 1)$ *e* $k = 0$.

Passo 2: *Encontrar um par* $(\tilde{x}^k, e^k) \in \mathbb{R}^n \times \mathbb{R}^n$ *tal que* $\tilde{x}^k$ *resolve EP* $(\tilde{L}^e_k, \mathbb{R}^n)$ *com* $(\tilde{L}^e_k$ definida como:

$$\tilde{\mathcal{L}}^e_k(x, y) = f(x, y) + \gamma_k\langle x - x^k, y - x\rangle + \sum_{i=1}^{m} p_i(x, y, \lambda^k, \gamma_k) - \langle e^k, y - x\rangle,$$

com p_i dada por (3.3) *e e^k dado por:*

$\|e^k\| \leq \delta\gamma_k \|\tilde{x}^k - x^k\|.$

Passo 3: *Atualize a variável dual*

$$\lambda_i^{k+1} = \max\left\{0, \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k}\right\}, \quad (i = 1, \ldots, m).$$

Passo 4: *Se* $(x^k, \lambda^k) = (\tilde{x}^k, \lambda^{k+1})$ *pare. Senão faça:*

$$x^{k+1} = \tilde{x}^k - \frac{1}{\gamma_k} e^k,$$

$k = k + 1$ *e volte ao Passo 2.*

Note que no Passo 2 do Algoritmo IALEM, a função lagrangeano aumentado regularizada apresenta o termo $\langle e^k, y - x\rangle$, o qual representa o erro. Daí a presença do termo inexato na denominação do algoritmo.

A seguir apresentamos o Algoritmo de Ponto Proximal Extragradiente Inexato, o qual será utilizado como resultado auxiliar.

Algoritmo 3.2 ***Ponto Proximal Extragradiente Inexato (IPPEM)***

Passo 1: *Considere uma sequência limitada* $\{\gamma_k\} \subset \mathbb{R}_{++}$, $x^0 \in K$, *uma tolerância* $\delta \in (0, 1)$ *e* $k = 0$.

Passo 2: *Encontre um par* $(\hat{x}^k, e^k) \in \mathbb{R}^n \times \mathbb{R}^n$ *tal que* $\hat{x}^k$ *resolve* $EP(f_k^e, K)$ *com*

$f_k^e(x, y) = f(x, y) + \gamma_k\langle x - x^k, y - x\rangle - \langle e^k, y - x\rangle$ *e* $\|e^k\| \leq \delta\gamma_k\|\hat{x}^k - x^k\|.$

Passo 3: *Se* $\hat{x}^k = x^k$, *então pare. Senão faça:*

$$x^{k+1} = \hat{x}^k - \frac{1}{\gamma_k} e^k,$$

$k = k+1$,

e volte ao Passo 2.

A análise de convergência do Algoritmo IPPEM é realizada usando as técnicas já conhecidas dos métodos de ponto proximal. Desta forma, (NASRI, 2010) mostram que o

Algoritmo IPPEM converge para um ponto $x^* \in S(f, K)$.

Já a prova de convergência do Algoritmo IALEM é realizada também pelos procedimentos padrões conhecidos na literatura de programação não linear, através da equivalência entre primal-dual. Para isso, são necessárias duas hipóteses adicionais:

P4: Considere a função f dada em $EP(f, K)$, θ sub monótona, isto é, existe $\theta \geq 0$ tal que $f(x, y) + f(y, x) \leq \theta \|x - y\|^2, \forall x, y \in \mathbb{R}^n$.

P5: Para todo $x^1, \ldots, x^q \in \mathbb{R}^n$ e todo $t_1, \ldots, t_q \mathbb{R}_+$ tal que $\sum_{i=1}^{q} t_i = 1$, se cumpre que

$$\sum_{i=1}^{q} t_i f\left(x^i, \sum_{k=1}^{q} t_k x^k\right) \leq 0.$$

No próximo teorema apresentamos a prova de convergência do Algoritmo IPPEM.

Teorema 3.1 *Teorema de convergência do IPPEM*

Considere $EP(f, K)$ *satisfazendo* ***P1-P5*** *e uma sequência exógena* $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$ *para algum* $\bar{\gamma} > \theta$*, sendo* θ *a constante de monotonicidade em* ***P4****. Se* $EP(f, K)$ *tem solução e* $\{x^k\}$ é a sequência gerada pelo Algoritmo IPPEM, então $\{x^k\}$ *converge para alguma solução* x^* *de* $EP(f, K)$ *.*

Prova. Ver Teorema 5.8 de (IUSEM; NASRI, 2007).

No Teorema 3.1, o problema de equilíbrio considerado é $EP(f, K)$ significando a dependência do conjunto K. Na próxima proposição esta dependência será relaxada, ou seja, o Algoritmo IPPEM será aplicado para resolver $EP(L, \mathbb{R}^n \times \mathbb{R}^m)$, com L dada em (3.2).

Proposição 3.1 *Considere f satistazendo* ***P1-P5*** *em* $\mathbb{R}^n \times \mathbb{R}^n$ *e K dado por* (3.1)*. Então* L*, como definida em* (3.2)*, satisfaz* ***P1-P5*** *em* $(\mathbb{R}^n \times \mathbb{R}^m_+) \times (\mathbb{R}^n \times \mathbb{R}^m_+)$.

Prova. Ver (NASRI, 2010)

No Passo 2 do Algoritmo IPPEM será necessário resolver o seguinte subproblema:

Obter um par $(\bar{x}^k, e^k) \in \mathbb{R}^n \times \mathbb{R}^n$ tal que $\bar{x}^k$ resolve $EP(f^e_k, K)$ sendo

$$f^e_k(x, y) = f(x, y) + \gamma_k \langle x - x^k, y - x \rangle - \langle e^k, y - x \rangle, \tag{3.5}$$

e o erro satisfazendo

$$\|e^k\| \leq \delta \gamma_k \|\bar{x}^k - x^k\|. \tag{3.6}$$

Poderemos, agora, aplicar o Algoritmo IPPEM, para resolver o problema de equilí-

brio irrestrito $EP(L, \mathbb{R}^n \times \mathbb{R}^m_+)$. Para isso, na iteração k deste algoritmo será necessário determinar a função de regularização L deste novo problema. Utilizando a relação (3.5) com L no lugar de f temos:

$$\hat{L}^e_k((x, \lambda), (y, \mu)) =$$

$$= L((x, \lambda), (y, \mu)) + \gamma_k\langle (x, \lambda) - (x^k, \lambda^k), (y, \mu) - (x, \lambda)\rangle - \langle e^k, (y, \mu) - (x, \lambda)\rangle =$$

$$= L((x, \lambda), (y, \mu)) + \gamma_k\langle x - x^k, \lambda - \lambda^k\rangle + \gamma_k\langle y - x, \mu - \lambda\rangle - \langle e^k, y - x\rangle. \qquad (3.7)$$

em que L é dada por (3.2) e não foi usado o erro associado a λ e μ no argumento de L. Isto está relacionado ao fato que no Passo 3 do Algoritmo IALEM a variável λ é atualizada por uma fórmula fechada e, desta forma, é assumido que a atualização é realizada de uma maneira exata.

O Algoritmo IPPEM para resolver $EP(L, \mathbb{R}^n \times \mathbb{R}^m_+)$ será denotado daqui em diante por IPPEMI, ou seja, Ponto Proximal Extragradiente Inexato Irrestrito, o qual é dado por:

Algoritmo 3.3 Ponto Proximal Extragradiente Inexato Irrestrito (IPPEMI)

Passo 1: *Considere uma sequência limitada* $\{\gamma_k\} \subset \mathbb{R}_{++}$, $(\hat{x}^0, \hat{\lambda}^0) \in \mathbb{R}^n \times \mathbb{R}^m_+$, $e^0 \in \mathbb{R}^n$, $k = 0$ *e uma tolerância* $\delta \in (0, 1)$.

Passo 2: *Encontre* $(\hat{x}^k, \hat{\lambda}^k) \in \mathbb{R}^n \times \mathbb{R}^m_+$ *e* $(e^k, 0) \in \mathbb{R}^n \times \mathbb{R}^m$ *tal que* $(\hat{x}^k, \hat{\lambda}^k)$ *resolve* $EP(\hat{L}^e_k, \mathbb{R}^n \times \mathbb{R}^m_+)$ *com*

$$\hat{L}^e_k((x, \lambda), (y, \mu)) = L((x, \lambda), (y, \mu)) + \gamma_k\langle x - x^k, y - x\rangle + \gamma_k\langle \lambda - \lambda^k, \mu - \lambda\rangle - \langle e^k, y - x\rangle.$$

e

$$\|e^k\| \leq \delta\gamma_k\|(\hat{x}^k - x^k, \hat{\lambda}^k - \lambda^k)\|.$$

Passo 3: *Se* $\hat{x}^k = x^k$ *e* $\hat{\lambda}^k = \lambda^k$, *então pare. Caso contrário faça*

$$x^{k+1} = \hat{x}^k - \frac{1}{\gamma_k}e^k \quad e \quad \lambda^{k+1} = \hat{\lambda}^k. \qquad (3.8)$$

O próximo resultado, apesar de tratar da convergência do Algoritmo IPPEMI, foi apresentado em (NASRI, 2010) como um corolário, em razão do Algoritmo IPPEMI ser uma decorrência do Algoritmo IPPEM. Lembrando que a função $\hat{L}^e_k$ e do Algoritmo IPPEMI, dada

na relação (3.7), foi obtida da função f_k^e do Algoritmo IPPEM, dada na relação (3.5).

Corolário 3.1 *Convergência do Algoritmo IPPEMI*

Considere EP (f, K) *com K dado por* (3.1), *f satisfazendo* ***P1-P5*** *em* $R^n \times R^n$ *e uma sequência* $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$ *para algum* $\bar{\gamma} > \theta$, *em que* θ *é a constante de sub monotonicidade de f. Seja* $\{(x^k, \lambda^k)\}$ *a sequência gerada pelo Algoritmo IPPEMI. Se o problema EP* $(L, R^n \times R^m_+)$ *tem solução, então* $\{(x^k, \lambda^k)\}$ *converge para algum par* $(x^*, \lambda^*) \in S(L, R^n \times R^m_+)$.

Prova. Ver (NASRI, 2010).

Antes de tratar da equivalência da sequência dual gerada pelos Algoritmos IPPEMI e IALEM será necessário estabelecer a associação entre o conjunto de soluções de *EP* (f, K) e de *EP* $(L, R^n \times R^m_+)$. Esta relação será útil para garantir a unicidade de solução no passo principal de IPPEMI. Para isso serão necessários alguns resultados preliminares, os quais descrevemos na sequência.

Para cada $x \in R^n$, define-se $F_x : R^n \to R$ como

$$F_x(y) = f(x, y). \tag{3.9}$$

a qual é convexa em decorrência da propriedade **P2**. Esta função será usada para estabelecer a relação entre $S(f, K)$ e $S(L, R^n \times R^m_+)$.

Proposição 3.2 *Considere EP* (f, K). *As duas afirmações seguintes são equivalentes:*

i) $x^* \in S(f, K)$.

ii) x^* *minimiza* F_{x^*} *em K, sendo* F_{x^*} *dada por* (3.9).

Prova. Suponha que $x^* \in S(f, K)$. De (3.9) e *P* 1 tem se

$$F_{x^*}(y) = f(x^*, y) \geq= f(x^*, x^*) = F_{x^*}(x^*)$$

para todo $y \in K$, estabelecendo (*ii*). Agora suponha que (*ii*) é satifeita. Usando novamente **P1** e (3.9), tem se

$$f(x^*, y) = F_{x^*}(y) \geq F_{x^*}(x^*) = f(x^*, x^*) = 0$$

para todo $y \in K$, o que fornece o resultado desejado.

Para os próximos resultados será necessário o conceito de par ótimo para o problema *EP* (f, K), ou seja:

Definição 3.3 *O par* $(x^*, \lambda^*) \in R^n \times R^m$ *será chamado um par ótimo para EP* (f, K) *se*

$$0 \in \partial F_{x^*}(x^*) + \sum_{i=1}^{m} \lambda_i^* \partial h_i(x^*), \qquad (3.10)$$

$$\lambda_i^* \geq 0 \quad (1 \leq i \leq m), \qquad (3.11)$$

$$h_i(x^*) \leq 0 \quad (1 \leq i \leq m), \qquad (3.12)$$

$$\lambda_i^* h_i(x^*) = 0 \quad (1 \leq i \leq m), \qquad (3.13)$$

em que $\partial F_{x^*}(x^*)$, $\partial h_i(x^*)$ denotam o subdiferencial das funções convexas F_{x^*} e h_i, respectivamente, no ponto x^*. Observe que (3.10) - (3.13) são as condições de KKT associadas ao problema

Minimizar $F_{x^*}(x)$ *sujeito a* $x \in K$,

sendo K dado por (3.1).

As próximas proposições são resultados auxiliares para se obter a equivalência entre as soluções de $EP(f, K)$ e de $EP(L, R^n \times R^m_+)$.

Proposição 3.3 *Considere EP* (f, K) *e assuma que f cumpre* ***P1-P3*** *em* $R^n \times R^n$*. Então as seguintes afirmações são equivalentes.*

i) (x^*, λ^*) é *um par* ótimo *de EP* (f, K).

ii) $(x^*, \lambda^*) \in S(L, R^n \times R^m_+)$.

Prova. Ver (NASRI, 2010)

Proposição 3.4 *Considere EP* (f, K) *e assuma que f cumpre* ***P1-P3*** *em* $R^n \times R^n$*. Se* x^* *minimiza* F_{x^*} *em K, com* F_{x^*} *como em* (3.9) *e a condição de Qualificação de Slater dada na definição 3.2 se verifica para as funções* h_i *que definem o conjunto K, então existe* $\lambda^* \in R^m_+$ *tal que* (x^*, λ^*) *é um par* ótimo *para EP* (f, K)*. Reciprocamente, se* (x^*, λ^*) *é um par* ótimo *para EP* (f, K)*, então* x^* *minimiza* F_{x^*} *em K, com* F_{x^*} *como em* (3.9).

Prova. Ver (NASRI, 2010)

Corolário 3.2 *Considere EP* (f, K) *e assuma que f cumpre* ***P1-P3*** *em* $R^n \times R^n$*. Se* $(x^*, \lambda^*) \in S(L, R^n \times R^m)$*, então* $x^* \in S(f, K)$*. Reciprocamente, se* $x^* \in S(f, K)$ *e a Condição de Qualificação de Slater dada na definição 3.2 se verifica, então existe* $\lambda^* \in R^m_+$ *tal que* $(x^*, \lambda^*) \in S(L, R^n \times R^m_+)$.

Prova. Ver (NASRI, 2010)

O Corolário 3.2 mostra que resolver $EP(L, R^n \times R^m_+)$ é suficiente para resolver *EP*

(f, K). Ou seja, este resultado garante que resolver o problema de equilíbrio irrestrito é suficiente para se resolver o problema de equilíbrio restrito.

O resultado mais importante deste capítulo será o teorema de equivalência, ou seja, o teorema que garante que a sequência gerada pelo Algoritmo IPPEMI para resolver *EP* $(L, R^n \times R^m_+)$ coincide com a sequência gerada pelo Algoritmo IALEM para resolver *EP* (f, K). Antes, porém, apresentaremos um resultado técnico que será utilizado na prova de equivalência.

Proposição 3.5 *Considere f satisfazendo* **P1-P4**, $z \in R^n$ *e* $\gamma > \theta$, *sendo* θ *a constante de sub monotonicidade de f dada em* **P4**. *Se* $\tilde{f} : K \times K \to R$ *é definida como*

$$\tilde{f}(x, y) = f(x, y) + \gamma\langle x - z, y - x\rangle - \langle e, y - x\rangle,$$

então EP $(\tilde{f}, K)$ *tem uma única solução.*

Prova. Ver (NASRI, 2010)

Com o que foi desenvolvido até aqui tem se todas as ferramentas necessárias para apresentar o teorema de equivalência. Apesar deste teorema estar demonstrado em (NASRI, 2010), exibiremos sua prova aqui também, por se tratar de um resultado importante na abordagem que estamos apresentado nesta tese.

Teorema 3.2 *Equivalência para o Problema de Equilíbrio*

Considere EP (f, K) *satisfazendo as propriedades* **P1-P4** *em* $R^n \times R^n$, *uma sequência* $\{\gamma_k\} \subset R_{++}$, *uma tolerância* $\delta \in (0, 1)$ *e* $\{(x^k, \lambda^k)\}$ *a sequência gerada pelo Algoritmo IALEM aplicado a EP* (f, K), *com o vetor erro associado* $e^k \in R^n$.

Seja $\{(\bar{x}^k, \bar{\lambda}^k)\}$ *a sequência gerada pelo Algoritmo IPPEMI aplicado a EP* $(L, R^n \times R^m)$, *com vetor erro* $(e^k, 0) \in R^n \times R^m$, *com os mesmos* γ_k *e* δ. *Se* $(x^0, \lambda^0) = (\bar{x}^0, \bar{\lambda}^0)$, *então* $\{(x^k, \lambda^k)\} = \{(\bar{x}^k, \bar{\lambda}^k)\}$, *para todo k.*

Prova. A prova será feita por indução sobre k. Para $k = 0$ o resultado é válido por hipótese. Suponha que $(x^k, \lambda^k) = (\bar{x}^k, \bar{\lambda}^k)$. Precisamos mostrar que

$$(x^{k+1}, \lambda^{k+1}) = (\bar{x}^{k+1}, \bar{\lambda}^{k+1}).$$

Em vista do Passo 2 do Algoritmo IPPEMI é necessário resolver *EP* $(L, R^n \times R^m_+)$, com L^e_k dada por (3.7), o qual tem solução única pela Proposição (3.5). Pela Proposição (3.4) $(\hat{x}^k, \hat{\lambda}^k)$ resolve o problema de minimização convexa definido como

$$Minimizar\ \hat{\mathcal{F}}_{(\hat{x}^k,\hat{\lambda}^k)}(x,\lambda)$$
$$s.a.\ (x,\lambda) \in \mathbb{R}^n \times \mathbb{R}^m_+. \tag{3.14}$$

com $F(\hat{x}^k, \hat{\lambda}^k)(x, \lambda) = \hat{L}^e_k((\hat{x}^k, \hat{\lambda}^k), (x, \lambda))$. As restrições deste problema são afim tal que a condição de Qualificação de Slater, dada na definição 3.2, se verifica e contudo existe um vetor KKT $u^k \in R^m$ (ver, por exemplo, Teorema 2.3.2 do Capítulo VII de (HIRIART- URRUTY; Lemaréchal, 1996), o qual trata com do caso não suave) tal que

$$[\gamma_k(\bar{x}^k - \hat{x}^k) + e^k] \in [\partial F_{\hat{x}^k}(\hat{x}^k) + \sum_{i=1}^{m} \hat{\lambda}_i^k \partial h_i(\hat{x}^k)], \tag{3.15}$$

$$-h_i(\hat{x}^k) + \gamma_k(\hat{x}_i^k - \bar{\lambda}_i^k) = u_i^k \quad (1 \leq i \leq m), \tag{3.16}$$

$$\hat{\lambda}^k \geq 0, \tag{3.17}$$

$$u^k \geq 0, \tag{3.18}$$

$$\hat{\lambda}_i^k u_i^k = 0 \quad (1 \leq i \leq m). \tag{3.19}$$

Observe que (3.15) e (3.16) são os subdiferenciais de $\hat{L}^e_k((\hat{x}^k, \hat{\lambda}^k), (x, \lambda))$ calculados com respeito a x e λ, respectivamente, no ponto $(x, \lambda) = (\hat{x}^k, \hat{\lambda}^k)$. Além disso, foi usado nestas equações, a hipótese de indução. Combinando (3.15) a (3.19) obtemos:

$$\hat{\lambda}_i^k = \max\left\{0, \bar{\lambda}_i^k + \frac{h_i(\hat{x}^k)}{\gamma_k}\right\} \quad (1 \leq i \leq m). \tag{3.20}$$

Substituindo (3.20) na equação (3.15) temos:

$$[\gamma_k(\bar{x}^k - \hat{x}^k) + e^k] \in \left[\partial F_{\hat{x}^k}(\hat{x}^k) + \sum_{i=1}^{m} \max\left\{0, \bar{\lambda}_i^k + \frac{h_i(\hat{x}^k)}{\gamma_k}\right\} \partial h_i(\hat{x}^k)\right]. \tag{3.21}$$

No Passo 2 do Algoritmo IALEM, precisamos obter a solução $\tilde{x}^k$ de $EP(\tilde{L}^e_k, R_n)$. Dessa forma, aplicando a Proposição 3.2 obtemos que $\tilde{x}^k \in EP(\tilde{L}^e_k, R_n)$ se, e somente se,

$$[\gamma_k(x^k - \tilde{x}^k) + e^k] \in \left[\partial F_{\tilde{x}^k}(\tilde{x}^k) + \sum_{i=1}^{m} \max\left\{0, \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k}\right\} \partial h_i(\tilde{x}^k)\right]. \tag{3.22}$$

Usando a hipótese de indução, ou seja, $x^k = \bar{x}^k$ e $\lambda^k = \bar{\lambda}^k$, obtemos de (3.21) que (3.22) se verifica com $\hat{x}^k$ no lugar de $\tilde{x}^k$ e, então, $\hat{x}^k$ também resolve $EP(\tilde{L}^e_k, R^n)$. Desde que, pela Proposição 3.5, este problema tem solução única, podemos concluir que

$$\hat{x}^k = \tilde{x}^k. \tag{3.23}$$

Pelo Passo 4 do Algoritmo IALEM, tem-se $x^{k+1} = \tilde{x}^k - \frac{1}{\gamma_k} e^k$ e, por outro lado, pelo Passo γ_k do Algoritmo IPPEMI, temos $\bar{x}^{k+1} = \hat{x}^k - \frac{1}{\gamma_k} e^k$, o que pode ser concluído, pela equação (3.23), que $x^{k+1} = \bar{x}^{k+1}$. Ainda falta a conclusão em relação à variável dual. Pelo Passo 3 do Algoritmo IPPEMI $\bar{\lambda}^{k+1} = \bar{\lambda}^k$. Pela equação (3.20)

$$\bar{\lambda}^{k+1} = \hat{\lambda}_i^k = \max\left\{0, \bar{\lambda}_i^k + \frac{h_i(\hat{x}^k)}{\gamma_k}\right\} \quad (1 \leq i \leq m). \tag{3.24}$$

(3.24)

Por outro lado, pelo Passo 3 do Algoritmo IALEM, tem se

$$\lambda^{k+1} = \max\left\{0, \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k}\right\} \quad (1 \leq i \leq m). \tag{3.25}$$

Utilizando a hipótese de indução, $\lambda^k = \bar{\lambda}^k$, em (3.25) juntamente com (3.23) obtemos que (3.24) é igual a (3.25), ou seja $\bar{\lambda}^{k+1} = \lambda^{k+1}$, completando o passo indutivo e a prova.

A próxima proposição trata da terminação finita do Algoritmo IALEM.

Proposição 3.6 *Suponha que o Algoritmo IALEM pare na iteração k. Então, o vetor* $\tilde{x}^k$ *gerado pelo algoritmo* é *uma solução de EP* (*f, K*).

Prova. Ver (NASRI, 2010)

Finalmente, com o uso do Teorema 3.2 a prova de convergência do Algoritmo IALEM pode ser completada, pelo seguinte teorema:

Teorema 3.3 *Considere EP* (*f, K*). *Suponha que*

*i) f satisfaça **P1-P5** em* $\mathbb{R}^n \times \mathbb{R}^n$,

ii) K é *dado por* (3.1),

iii) A Condição de Qualificação de Slater dada na definição 3.2 se verifica para K e

iv) $\gamma_k \subset (\theta, \bar{\lambda}]$ *para algum* $\bar{\lambda} > \theta$, *sendo* θ *a constante de sub monotonicidade de f dada em **P4**.*

Considere $\{(x^k, \lambda^k)\}$ *a sequência gerada pelo Algoritmo IALEM para resolver EP* (*f, K*). *Se EP* (*f, K*) *tem solução, então a sequência* $\{(x^k, \lambda^k)\}$ *converge para algum par* ótimo

(x^*, λ^*) *para* $EP(f, K)$ *e, consequentemente,* $x^* \in S(f, K)$.

Prova. Ver (NASRI, 2010)

É importante notar que a maior dificuldade a ser superada na implementação do Algoritmo IALEM é determinar a solução do subproblema gerado no Passo 2. Da demonstração do Teorema de Equivalência 3.2 ficou evidente que $\tilde{x}^k \in S(L_k^e, R^n)$ se, e somente se, a equação (3.22) é satisfeita, ou seja:

$$e^k \in \left[\partial F_{\tilde{x}^k}(\tilde{x}^k) + \sum_{i=1}^{m} \max\left\{0, \lambda_i^k + \frac{h_i(\tilde{x}^k)}{\gamma_k}\right\} \partial h_i(\tilde{x}^k)\right] + \gamma_k(\tilde{x}^k - x^k), \qquad (3.26)$$

na qual $\partial F\tilde{x}^k(\tilde{x}^k)$ é o subgradiente de F_x calculado em $\tilde{x}^k$, com F_x dada por (3.9), $\partial h_i(\tilde{x}^k)$ é o subgradiente de h_i calculado em $\tilde{x}^k$. Contudo, a Proposição 3.6 diz que se existe um procedimento que converge para a solução do problema de equilíbrio irrestrito $EP(\tilde{L}_k^e, R^n)$ com $e^k = 0$. Então tal procedimento, após um número finito de passos, irá gerar um par $(\tilde{x}^k, e^k)$ satisfazendo a condição do Passo 2 daquele algoritmo.

Cabe agora uma explicação do real significado do vetor erro e^k nos Algoritmos IALEM, IPPEM e IPPEMI, detalhada em (NASRI, 2010) e resumida a seguir.

Suponha que temos uma sub-rotina que é usada para resolver o k-ésimo subproblema gerado no Passo 2 do Algoritmo IALEM com $e^k = 0$. Neste cenário, a sub-rotina gera soluções aproximadas $\tilde{x}^{k,l}$ cujos vetores erros são $e^{k,l}$ para $l = 1, 2,$ Uma vez que $\tilde{x}^{k,l}$ é calculado em cada iteração l da sub-rotina, esta verifica se o vetor erro resultante $e^{k,l}$ satisfaz o critério dado no Passo 2. Se assim for, o par $(\tilde{x}^{k,l}, e^{k,l})$ é aceito e o algoritmo inicia uma nova iteração $k = k + 1$. Caso contrário, o algoritmo executa a sub-rotina para mais uma iteração e define $l = l + 1$. Em particular, considere o caso quando f e h_i são diferenciáveis. Nesta situação, usando (3.26), $\tilde{x}^{k,l}$ satisfaz

$$e^{k,\ell} = \gamma_k(\tilde{x}^{k,\ell} - x^k) + \nabla F_{\tilde{x}^{k,\ell}}(\tilde{x}^{k,\ell}) + \sum_{i=1}^{m} \max\left\{0, \lambda_i^k + \frac{h_i(\tilde{x}^{k,\ell})}{\gamma_k}\right\} \nabla h_i(\tilde{x}^{k,\ell}). \qquad (3.27)$$

Se $\tilde{x}^{k,l}$ é a solução exata do k-ésimo subproblema, então o lado direito de (3.27) desaparece. Caso contrário, $\tilde{x}^{k,ll}$ aproxima a solução exata e o lado direito de (3.27) é diferente de zero e é chamado de $e^{k,l}$. Agora, se $e^{k,l}$ satisfaz a desigualdade do erro requerida no Passo 2 do Algoritmo IALEM com $\tilde{x}^k = \bar{x}^{k,l,l}$, então consideramos $e^k = e^{k,l}$ e executamos o Passo 3 do algoritmo. Se $\tilde{x}^{k,l}$ proposto não é bom o suficiente, é necessário um passo adicional da sub-rotina auxiliar, após o qual o teste será repetido com $\tilde{x}^k = \tilde{x}^{k,l+1}$ e $e^k = e^{k,l+1}$. Pela Proposição (3.6), após um número finito de iterações, um par adequado $(\tilde{x}^k, e^k)$ deve ser gerado satisfazendo a desigualdade do erro requerida no Passo 2 deste algoritmo.

No caso suave, cada iteração k do Algoritmo IALEM consiste em resolver uma equação não linear $G(x) = 0$, em que $G : R^n \to R^n$ é dada por

$$G^k(x) = \gamma_k(x - x^k) + \nabla F_x(x) + \sum_{i=1}^{m} \max\left\{0, \lambda_i^k + \frac{h_i(x)}{\gamma_k}\right\} \nabla h_i(x). \qquad (3.28)$$

Observamos que o lado direito de (3.28) é contínuo, porém não diferenciável, devido à presença do máximo. Em (FERREIRA, 2013) foram propostas duas metodologias para resolver o problema de equilíbrio irrestrito utilizando a relação (3.28) ao invés do Passo 2 do Algoritmo IALEM. Uma das abordagens utilizadas substituiu o termo não diferenciável por uma aproximação suave que o tornou diferenciável e, então, o método de Newton foi utilizado para resolver o sistema suave gerado. A outra abordagem consistiu na utilização de uma família de métodos subgradientes para minimizar $\|G^k(x)\|^2$ com G^k dada por (3.28).

Dando continuidade aos trabalhos desenvolvidos por (FERREIRA, 2013) e (NASRI, 2010), esta pesquisa tem como um dos principais objetivos a resolução do problema de equilíbrio utilizando o método de lagrangeano aumentado, porém com penalidade diferente daquela usada em (NASRI, 2010). Desta forma, o subproblema gerado no Passo 2 do Algoritmo IALEM, será equivalente a um sistema não linear do tipo (3.28), e então, podemos utilizar métodos de otimização diferenciável para resolvê-lo.

LAGRANGEANO AUMENTADO EXPONENCIAL PARA O PROBLEMA DE EQUILÍBRIO

Neste capítulo apresentaremos a principal contribuição desta tese, ou seja, o Método de Lagrangeano Aumentado Extragradiente Exponencial Inexato aplicado ao Problema de Equilíbrio *EP* (*f, K*) dado na Definição 3.1. O método de lagrangeano aumentado exponencial, aplicado ao problema de otimização, foi descrito no Capítulo 2 e agora será estendido para o Problema de Equilíbrio Geral, como desenvolvido em (NASRI, 2010) e (FERREIRA, 2013). Além disso, abordaremos aspectos teóricos e computacionais acerca do algoritmo proposto.

1 | DESCRIÇÃO DOS ALGORITMOS

Nesta seção, apresentaremos o algoritmo de lagrangeano aumentado extragradiente exponencial inexato, assim como o algoritmo de ponto proximal extragradiente inexato.

Considere $p_i : IR^n \times IR^n \times IR^m \times IR_{++} \to IR$, para $i = 1, 2, \ldots, m$, e a função lagrangeano aumentado $\bar{L} : IR^n \times IR^n \times IR^m \times IR^n \times IR_{++} \to IR$ dadas, respectivamente, por

$$p_i(x, y, \lambda, \gamma) = \gamma \left[\lambda_i e^{\frac{1}{\gamma}h_i(y)} - \lambda_i e^{\frac{1}{\gamma}h_i(x)}\right] \tag{4.1}$$

e

$$\bar{L}(x, y, \lambda, z, \gamma) = f(x, y) + \gamma\langle x - z, y - x\rangle + \sum_{i=1}^{m} p_i(x, y, \lambda, \gamma). \tag{4.2}$$

Neste caso, o Algoritmo IALEM 3.1 será reescrito incluindo estas novas funções e será chamado de IALEMe (o último **e** é de exponencial).

Algoritmo 4.1 Lagrangeano Aumentado Extragradiente Exponencial Inexato (IALEMe)

Passo 1: *Considere uma sequência limitada* $\{\gamma_k\} \subset \mathbb{R}_{++}$, $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m_+$, *uma tolerância* $\delta \in (0, 1)$ *e* $k = 0$.

Passo 2: *Encontrar um par* (x^k, e^k) *tal que* $\tilde{x}^k$ *resolve EP* $(\tilde{L}^e_k, \mathbb{R}^m)$ *com* $\tilde{L}^e_k$ *definida como:*

$$\tilde{L}^e_k(x, y) = f(x, y) + \gamma_k\langle x - x^k, y - x\rangle + \sum_{i=1}^{m} p_i(x, y, \lambda^k, \gamma_k) - \langle e^k, y - x\rangle ,$$

com p_i *dada por* (4.1) *e* e^k *dado por:*

$\|e^k\| \leq \delta\gamma_k\|\tilde{x}^k - x^k\|$.

Passo 3: *Atualize a variável dual*

$$\lambda_i^{k+1} = \lambda_i^k e^{\frac{1}{\gamma_k} h_i(\tilde{x})}, \quad (i = 1, \ldots, m).$$

Passo 4: *Se* $(x^k, \lambda^k) = (\bar{x}^k, \lambda^{k+1})$, *então pare. Senão faça:*

$$x^{k+1} = \tilde{x}^k - \frac{1}{\gamma_k} e^k$$

$k = k + 1$

e volte ao Passo 2.

Da mesma forma o Algoritmo IPPEMI 3.3 será reescrito e será denominado IPPEMIe.

NOTA: será usada a notação vetorial para representar o logaritmo natural de $u \in \mathbb{R}^m_{++}$, ou seja, $\ln(u) = (\ln(u_1), \ldots, \ln(u_m))^T$.

Algoritmo 4.2 ***Ponto Proximal Extragradiente Inexato (IPPEMIe)***

Passo 1: *Considere uma sequência limitada* $\{\gamma_k\} \subset \mathbb{R}_{++}$, $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m_{++}$, $e^0 \in I\!R^n$, *uma tolerância* $\delta \in (0, 1)$ *e* $k = 0$.

Passo 2: *Encontre* $(\hat{x}^k, \hat{\lambda}^k) \in \mathbb{R}^n \times \mathbb{R}^m_{++}$ *e* $(e^k, 0) \in I\!R^n \times I\!R^m$ *tal que* $(\hat{x}^k, \hat{\lambda}^k)$ *resolve EP* $(\hat{L}^e_k, I\!R^n \times I\!R^m)$ *com*

$$\hat{L}^e_k((x, \lambda), (y, \mu)) = L((x, \lambda), (y, \mu)) + \gamma_k\langle x - x^k, y - x\rangle + \gamma_k\langle \ln(\lambda) - \ln(\lambda^k), \mu - \lambda\rangle - \langle e^k, y - x\rangle, \tag{4.3}$$

L *dada por* (3.2) *e* e^k *satisfazendo*

$$\|(e^k, 0)\| = \|e^k\| \leq \delta\gamma_k\|(\hat{x}^k - x^k, \hat{\lambda}^k - \lambda^k)\|.$$

Passo 3: *Se* $(\hat{x}^k, \lambda^k) = (x^k, \hat{\lambda}^k)$, *então pare. Senão faça:*

$$x^{k+1} = \hat{x}^k - \frac{1}{\gamma_k} e^k \quad e \quad \lambda^{k+1} = \hat{\lambda}^k$$

$k = k + 1$ *e volte ao Passo 2.*

Observe que a diferença fundamental entre os Algoritmos IPPEMIe e IPPEMI está na função que regulariza o problema. Na função (4.3), o termo de regularização é

$\gamma_k\langle \ln(\lambda) - \ln(\lambda^k), \mu - \lambda\rangle$ ao invés de $\gamma_k\langle \lambda - \lambda^k, \mu - \lambda\rangle$. Com esta função é possível mostrar a convergência do Algoritmo IPPEMIe, assim como a equivalência (dual) entre os Algoritmos IPPEMIe e IALEMe.

A seguir apresentaremos a convergência do Algoritmo IPPEMIe. Estamos supondo que a boa definição do Algoritmo IPPEMIe é decorrente da boa definição do Algoritmo IPPEMI, tomando-se os devidos cuidados com o domínio da variável dual λ em ln(λ), ou seja, λ deve ser não negativo.

2 | CONVERGÊNCIA DO ALGORITMO PROPOSTO

Nesta seção, nosso objetivo é tratar de aspectos teóricos relativos à convergência do Algoritmo 4.2, que estamos propondo. Conjecturamos que o mesmo é convergente. Porém, mostramos apenas uma espécie de convergência local, ou seja, conseguida impondo uma condição sobre o vetor de multiplicadores de Lagrange.

O resultado seguinte é clássico considerando a literatura de análise convexa, ver por exemplo (SOLODOV, 2005).

Teorema 4.1 *Considere $D \subset \mathbb{R}^n$ um conjunto aberto e convexo e $f : D \rightarrow \mathbb{R}$ uma função diferenciável em D. Então as propriedades seguintes são equivalentes:*

a. *A função f é convexa em D.*

b. *Para todos $x, y \in D$,*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle.$$

c. *Para todos $x, y \in D$,*

$$\langle \nabla f(y) - \nabla f(x), y - x\rangle \geq 0.$$

Quando f é duas vezes diferenciável em D, as propriedades acima também são equivalentes à

d. *A matriz Hessiana de f é semidefinida positiva em todo ponto de D, ou seja*

$$\langle \nabla^2 f(x)d, d\rangle \geq 0, \forall x \in D, \forall d \in \mathbb{R}^n.$$

Prova. Ver (SOLODOV, 2005).

Destacamos que no caso da função ser côncava as desigualdades acima continuam verdadeiras mas com sinais trocados. Por exemplo, no item (*a*) teremos, para todo $x, y \in D$,

$$f(y) \leq f(x) + \langle \nabla f(x), y - x\rangle.$$

Lema 4.1 *Considere $f : \mathbb{R}_{++} \rightarrow \mathbb{R}$ a função logaritmo natural $f(x) = \ln(x)$. Então*

(i) f é uma função côncava.

(ii) $\ln(x) \leq x - 1, \forall x \in \mathbb{R}_{++}$.

Prova. A prova do item (*i*) é imediata. Basta derivar duas vezes a função *f* e usar o item (*d*) do Teorema 4.1. Para o item (*ii*), considere o item (*b*) do Teorema 4.1 juntamente com o item (*i*) deste Lema, ou seja, $\ln(x) \leq \ln(y) + \frac{1}{y}(x - y)$. Como a desigualdade vale para todo $x, y \in \mathbb{R}_{++}$ basta escolher $y = 1$ e ter-se-à completado a prova.

Lema 4.2 *Considere* $u \in \mathbb{R}_{++}$ *e* $v, w \in \mathbb{R}$. *Então*

$$(1 + \ln(u) - u)(w - v) \geq 0 \Leftrightarrow w \leq v.$$

Prova. Desde que $\ln(u)$ é uma função côncava, pelo item (*ii*) do Lema 4.1

$$\ln(u) \leq u - 1 \Leftrightarrow 1 + \ln(u) - u \leq 0.$$

Consequentemente,

$$(1 + \ln(u) - u)(w - v) \geq 0 \Leftrightarrow w - v \leq 0 \Leftrightarrow w \leq v, \forall w \in \mathbb{R}.$$

Teorema 4.2 *Considere* $u, v \in \mathbb{R}^m_{++}$ *e* $w \in \mathbb{R}^m$. *Então,*

$$\langle v - u - \ln(v) + \ln(u), w - v\rangle \geq 0 \Leftrightarrow w \leq v.$$

Prova. $\langle v - u - \ln(v) + \ln(u), w - v\rangle = \sum_{i=1}^{m} [(v_i - u_i) + (\ln(u_i) - \ln(v_i))](w_i - v_i) \geq \sum_{i=1}^{m} (1 + \ln(u_i) - u_i)(w_i - v_i)$. Na última desigualdade foi usado o item (*ii*) do Lema 4.1, ou seja $\ln(v_i) \leq v_i - 1 \Leftrightarrow -\ln(v_i) + v_i \geq 1, \forall i = 1, \ldots, m$. Pelo Lema 4.2, para cada $i = 1, \ldots, m$, $(1 + \ln(u_i) - u_i)(w_i - v_i) \geq 0 \Leftrightarrow w_i \leq v_i, \forall w_i \in \mathbb{R}$ e portanto a prova está completa.

A seguir provaremos a convergência do Algoritmo IPPEMIe, no sentido do que foi explicado no início da seção, prova esta que será baseada na convergência do Algoritmo IPPEMI.

Teorema 4.3 *Convergência do Algoritmo IPPEMIe*

Considere EP (f, K) *com K dado por* (3.1), *f satisfazendo* ***P1-P5*** *em* $\mathbb{R}^n \times \mathbb{R}^n$ *e uma sequência* $\{\gamma_k\} \subset (\theta, \bar{\gamma}]$ *para algum* $\bar{\gamma} > \theta$, *em que* θ *é a constante de sub monotonicidade de f. Seja* $\{(\hat{x}^k, \hat{\lambda}^k)\}$ *a sequência gerada pelo Algoritmo IPPEMIe com* $\hat{\lambda}^k \geq \mu > 0$. *Se o problema EP* $(L, \mathbb{R}^n \times \mathbb{R}^m_{++})$ *tem solução*, *então* $\{(\hat{x}^k, \hat{\lambda}^k)\}$ *converge para algum par* $(x^*, \lambda^*) \in S(L, \mathbb{R}^n \times \mathbb{R}^m_{++})$.

Prova. Considere as sequências $\{(\hat{x}^k, \hat{\lambda}^k)\} \subset R^n \times R^m_{++}$ e $\{e^k\} \subset R^n$ geradas pelo Algoritmo IPPEMIe, 4.2, então pelo Passo 2 tem-se

$$\hat{L}^e_k((\hat{x}^k, \hat{\lambda}^k), (y, \mu)) \geq 0 \tag{4.4}$$

e

$$\|e^k\| \leq \delta\gamma_k\|(\hat{x}^k - x^k, \hat{\lambda}^k - \lambda^k)\|. \tag{4.5}$$

em que

$$\hat{L}^e_k((\hat{x}^k, \hat{\lambda}^k), (y, \mu)) = L((\hat{x}^k, \hat{\lambda}^k), (y, \mu)) + \gamma_k\langle \hat{x}^k - x^k, y - \hat{x}^k\rangle + \\ +\gamma_k\langle \ln(\hat{\lambda}^k) - \ln(\lambda^k), \mu - \hat{\lambda}^k\rangle - \langle e^k, y - \hat{x}^k\rangle, \tag{4.6}$$

e L é dada por (3.2).

Além disso, pelo Passo 3 do Algoritmo IPPEMIe tem-se

$$x^{k+1} = \hat{x}^k - \frac{1}{\gamma_k}e^k \quad \text{e} \quad \lambda^{k+1} = \hat{\lambda}^k. \tag{4.7}$$

Por outro lado, substituindo estes valores na função $\hat{L}^e_k$ e também na fórmula do erro do Passo 2 do Algoritmo IPPEMI, 3.3, tem-se

$$\hat{\mathcal{L}}^e_k((\hat{x}^k, \hat{\lambda}^k), (y, \mu)) = \mathcal{L}((\hat{x}^k, \hat{\lambda}^k), (y, \mu)) + \gamma_k\langle \hat{x}^k - x^k, y - \hat{x}^k\rangle + \gamma_k\langle \hat{\lambda}^k - \lambda^k, \mu - \hat{\lambda}^k\rangle - \langle e^k, y - \hat{x}^k\rangle, \tag{4.8}$$

com L dada por (3.2). Já a fórmula do erro coincide com (4.5) e, da mesma forma, a atualização de x e λ no Passo 3 de IALEMI coincide com (4.7). Agora, subtraindo (4.6) de (4.8) temos:

$$\hat{L}^e_k((\hat{x}^k, \hat{\lambda}^k), (y, \mu)) - \hat{L}^e_k((\hat{x}^k, \hat{\lambda}^k), (y, \mu)) = \gamma_k\langle \ln(\hat{\lambda}^k) - \ln(\lambda^k), \mu - \hat{\lambda}^k\rangle - \gamma_k\langle \hat{\lambda}^k - \lambda^k, \mu - \hat{\lambda}^k\rangle.$$

Isolando $\hat{L}^e_k$ no primeiro membro desta última igualdade e usando (4.4) tem-se

$$\hat{L}^e_k((\hat{x}^k, \hat{\lambda}^k), (y, \mu)) \geq \gamma_k\langle \hat{x}^k - \lambda^k, \mu - \hat{\lambda}^k\rangle - \gamma_k\langle \ln(\hat{\lambda}^k) - \ln(\lambda^k), \mu - \hat{\lambda}^k\rangle$$

ou, equivalentemente

$$\hat{L}^e_k((\hat{x}^k, \hat{\lambda}^k), (y, \mu)) \geq \gamma_k\langle \hat{\lambda}^k - \lambda^k - \ln(\hat{\lambda}^k) + \ln(\lambda^k), \mu - \hat{\lambda}^k\rangle. \tag{4.9}$$

Por hipótese $\mu \leq \hat{\lambda}^k$, então pelo Teorema 4.2 tem-se que

$$\langle \hat{\lambda}^k - \lambda^k - \ln(\hat{\lambda}^k) + \ln(\lambda^k), \mu - \hat{\lambda}^k\rangle \geq 0.$$

Desde que $\gamma_k > 0$, então (4.9) se torna.

$$\hat{L}^e_k((\hat{x}^k, \hat{\lambda}^k), (y, \mu)) \geq 0 \tag{4.10}$$

Logo, de (4.10), (4.5) e (4.7) tem-se todos os passos do Algoritmo IPPEMI satisfeitos e, portanto, pelo Corolário 3.1 se *EP* (L, $R^n \times R^m_{++}$) tem solução então $\{(\hat{x}^k, \hat{\lambda}^k)\}$ converge para algum par $(x^*, \lambda^*) \in S(L, R^n \times R^m_{++})$ e a prova está completa.

3 I EQUIVALÊNCIA ENTRE IALEM*E* E IPPEMI*E*

Nesta seção, mostraremos a equivalência entre as sequências geradas pelos Algoritmos IALEMe e IPPEMIe e, desta forma, obtemos a convergência do Algoritmo IALEMe.

Teorema 4.4 *Equivalência entre IALEMe e IPPEMIe Considere EP* (*f, K*) *satisfazendo* ***P1 - P4*** *em* $R^n \times R^n$*, uma sequência* $\{\gamma_k\} \in R_{++}$*, um erro relativo de tolerância* $\delta \in (0, 1)$ *e* $\{(x^k, \lambda^k)\}$ *a sequência gerada pelo Algoritmo IALEMe aplicado a EP* (*f, K*)*, com vetor erro associado* $e^k \in R^n$*. Por outro lado, considere* $\{(\bar{x}^k, \bar{\lambda}^k)\}$ *a sequência gerada pelo Algoritmo IPPEMIe aplicado ao EP* (L, $R^n \times R^m_{++}$), *com* L *dada por* (3.2) *e vetor erro associado* $(e^k, 0) \in R^n \times R^n$*, usando os mesmos* γ^k *e* δ*. Se* $(x^0, \lambda^0) = (\bar{x}^0, \bar{\lambda}^0)$*, então* $\{(x^k, \lambda^k)\} = \{(\bar{x}^k, \bar{\lambda}^k)\}$ *para todo k.*

Prova. A prova é feita por indução sobre k. Para $k = 0$, o resultado é verdadeiro por hipótese. Suponha que $(x^k, \lambda^k) = (\bar{x}^k, \bar{\lambda}^k)$ (Hipótese de indução). Provaremos que

$$(x^{k+1}, \lambda^{k+1}) = (\bar{x}^{k+1}, \bar{\lambda}^{k+1}).$$

Em vista do Passo 2 do Algoritmo IPPEMIe é necessário resolver *EP* (L^e_k, $R^n \times R^m$) com $L^e_k : R^n \times R_{++} \to R$ dada por (4.3). Pela Proposição (3.4), $\{(\hat{x}^k, \hat{\lambda}^k)\}$ resolve o problema de minimização convexa definido como

$$\begin{aligned} &\textit{Minimizar}\ \hat{F}(\hat{x}^k, \hat{\lambda}^k)(x, \lambda) \\ &s.a.\ (x, \lambda) \in R^n \times R^m_+. \end{aligned} \tag{4.11}$$

com $\hat{F}(\hat{x}^k, \hat{\lambda}^k)(x, \lambda) = L^e_k((\hat{x}^k, \hat{\lambda}^k), (x, \lambda))$, sendo L^e_k dada por (4.3). Observe que,

$$\partial_x \hat{F}(\hat{x}^k, \hat{\lambda}^k)(x, \lambda) = \partial_x L^e_k((x^k, \lambda^k), (x, \lambda)) = \partial_x L(\hat{x}^k, \hat{\lambda}^k), (x, \lambda) + \gamma_k(\hat{x}^k - x^k) - e^k,$$

com $L((\hat{x}^k, \hat{\lambda}^k), (x, \lambda)) = f(\hat{x}^k, x) + \sum_{i=1}^m \hat{\lambda}^k_i h_i(x) - \sum_{i=1}^m \lambda_i h_i(\hat{x}^k)$ (ver relação (3.2)) e

$$\partial_x L((\hat{x}^k, \hat{\lambda}^k)(x, \lambda) = \partial f(\hat{x}^k, x) + \sum_{i=1}^m \hat{\lambda}^k_i \partial h_i(x).$$

No ponto $x = \hat{x}^k$ e usando $F_x(y) = f(x, y)$ esta última equação resulta

$$\partial_x \mathcal{L}((\hat{x}^k, \hat{\lambda}^k), (x^k, \lambda)) = \partial F_{\hat{x}^k}(\hat{x}^k) + \sum_{i=1}^{m} \hat{\lambda}_i^k \partial h_i(\hat{x}^k). \text{ Logo, } 0 \in \partial_x \hat{\mathcal{F}}_{(\hat{x}^k, \hat{\lambda}^k)}(x^k, \lambda) \Leftrightarrow$$

$$0 \in \left[\gamma_k(\hat{x}^k - x^k) - e^k + \partial F_{\hat{x}^k}(\hat{x}^k) + \sum_{i=1}^{m} \hat{\lambda}_i^k \partial h_i(\hat{x}^k)\right] \Leftrightarrow$$

$$\Leftrightarrow \left[\gamma_k(x^k - \hat{x}^k) + e^k\right] \in \left(\partial F_{\hat{x}^k}(\hat{x}^k) + \sum_{i=1}^{m} \hat{\lambda}_i^k \partial h_i(\hat{x}^k)\right).$$

Estas equações serão utilizadas nas condições KKT para o problema (4.11). Desde que as restrições do problema (4.11) sejam afins, a Condição de Qualificação de Slater dada na Definição 3.2 se verifica. Contudo existe um vetor de KKT $v^k \in R^n$ (ver, por exemplo, Teorema 2.3.2 do Capítulo VII de (HIRIART-URRUTY; Lemaréchal, 1996), o qual trata com o caso não suave), tal que

$$[\gamma_k(\bar{x} - \hat{x}^k) + e^k] \in \left(\partial F_{\hat{x}^k}(\hat{x}^k) + \sum_{i=1}^{m} \hat{\lambda}_i^k \partial h_i(\hat{x}^k)\right) \tag{4.12}$$

$$-h_i(\hat{x}^k) + \gamma_k(\ln \hat{\lambda}_i^k - \ln \bar{\lambda}_i^k) = v_i^k, \qquad i = 1, \dots, m \tag{4.13}$$

$$\hat{\lambda}_i^k \geq 0, \quad i = 1, \dots, m \tag{4.14}$$

$$v_i^k \geq 0, \quad i = 1, \dots, m \tag{4.15}$$

$$\hat{\lambda}_i^k v_i^k = 0, \ i = 1, \dots, m \tag{4.16}$$

Observe que (4.12) e (4.13) são os subdiferenciais de $\hat{L}_k^e((\hat{x}^k, \hat{\lambda}^k), (x, \lambda))$ calculados com respeito a x e λ, respectivamente, no ponto $(x, \lambda) = (\hat{x}^k, \hat{\lambda}^k)$. Além disso, usamos nestas equações, a hipótese de indução.

De (4.13) - (4.16) temos

$$\gamma_k[\ln \hat{\lambda}_i^k - \ln \bar{\lambda}_i^k] = h_i(\hat{x}^k) \Leftrightarrow \ln(\hat{\lambda}_i^k / \bar{\lambda}_i^k) = \frac{1}{\gamma_k} \hat{h}_i(x^k)$$

ou seja,

$$\hat{\lambda}_i^k = \bar{\lambda}_i^k e^{\frac{1}{\gamma_k} h_i(\hat{x}^k)}. \tag{4.17}$$

Substituindo (4.17) em (4.12), obtemos

$$\left[\gamma_k(\bar{x} - \hat{x}^k) + e^k\right] \in \left[\partial F_{\hat{x}^k}(\hat{x}^k) + \sum_{i=1}^{m} \bar{\lambda}_i^k e^{\frac{1}{\gamma_k} \hat{h}_i(\hat{x}^k)} \partial h_i(\hat{x}^k)\right], \quad i = 1, \dots, m. \tag{4.18}$$

No Passo 2 do Algoritmo IALEMe precisamos determinar a solução $\tilde{x}^k$ de $EP(\tilde{L}_k^e, \mathbb{R}^n)$ com $\tilde{L}_k^e : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ dada por

$$\tilde{L}_k^e(x,y) = f(x,y) + \gamma_k\langle x - x^k, y - x\rangle + \sum_{i=1}^{m}\left[\lambda_i^k e^{\frac{1}{\gamma_k}h_i(y)} - \lambda_i^k e^{\frac{1}{\gamma_k}h_i(x)}\right] - \langle e^k, y - x\rangle.$$

Para isso, aplicando a Proposição 3.2 a este problema temos que $\tilde{x}^k \in S(\tilde{L}_k^e, \mathbb{R}^n)$ se, e somente se,

$$\left[\gamma_k(x^k - \tilde{x}^k) + e^k\right] \in \left[\partial F_{\hat{x}^k}(\tilde{x}^k) + \sum_{i=1}^{m}\lambda_i^k e^{\frac{1}{\gamma_i^k}h_i(\tilde{x}^k)}\partial h_i(\tilde{x}^k)\right]. \tag{4.19}$$

Pela hipótese de indução $x^k = \bar{x}^k$ e $\lambda^k = \bar{\lambda}^k$, então (4.19) torna-se

$$\left[\gamma_k(\bar{x}^k - \tilde{x}^k) + e^k\right] \in \left[\partial F_{\hat{x}^k}(\tilde{x}^k) + \sum_{i=1}^{m}\bar{\lambda}_i^k e^{\frac{1}{\gamma_i^k}h_i(\tilde{x}^k)}\partial h_i(\tilde{x}^k)\right]. \tag{4.20}$$

Desde de que, pela Proposição 3.5 este problema tem solução única, então (4.20) coincide com (4.18) quando

$$\tilde{x}^k = \hat{x}^k. \tag{4.21}$$

Pela atualização de x, no Passo 4 do Algoritmo IALEMe,

$$x^{k+1} = \tilde{x}^k - \frac{1}{\gamma_k}e^k \tag{4.22}$$

e pela atualização de x, no Passo 3 do Algoritmo IPPEMIe,

$$\bar{x}^{k+1} = \tilde{x}^k - \frac{1}{\gamma_k}e^k. \tag{4.23}$$

De (4.21), (4.22) e (4.23), concluimos que $x^{k+1} = \bar{x}^{k+1}$.

Pela atualização das variáveis duais no Passo 3 do Algoritmo IPPEMIe e por (4.17), temos:

$$\bar{\lambda}_i^{k+1} = \bar{\lambda}_i^k e^{\frac{1}{\gamma_k}h_i(\hat{x}^k)}. \tag{4.24}$$

Agora, comparando (4.24) com $\lambda_i^{k+1} = \bar{\lambda}_i^k e^{\frac{1}{\gamma_k}h_i(\hat{x}^k)}$, dado no Passo 3 do Algoritmo 4.1, juntamente com a hipótese de indução, concluímos que $\lambda^{k+1} = \bar{\lambda}^{k+1}$, completando o passo indutivo e a prova.

4 I AJUSTE QUADRÁTICO DA PENALIDADE EXPONENCIAL

Na seção 3.2, em que foi considerado o caso suave, a cada iteração k do Algoritmo IALEM era necessário a resolução de um sistema não linear $G(x) = 0$, em que $G : \mathbb{R}^n \to \mathbb{R}^n$ é dada por

$$G^k(x) = \gamma_k(x - x^k) + \nabla F_x(x) + \sum_{i=1}^{m} \max\left\{0, \lambda_i^k + \frac{h_i(x)}{\gamma_k}\right\} \nabla h_i(x). \qquad (4.25)$$

A principal diferença entre os Algoritmos IALEM e IALEMe é uso da penalidade exponencial no lugar da penalidade quadrática clássica. Desta forma, com relação ao Algoritmo IALEMe, o sistema não linear 4.25 pode ser reescrito como

$$G^k(x) = \gamma_k(x - x^k) + \nabla F_x(x) + \sum_{i=1}^{m} \lambda_i^k e^{\frac{1}{\gamma_k} h_i(x)} \nabla h_i(x). \qquad (4.26)$$

Em geral, os métodos de lagrangeanos aumentados exponenciais tendem a gerar subproblemas mau condicionados devido ao rápido crescimento da função penalidade exponencial. Uma alternativa para contornar o mau condicionamento é fazer um ajuste quadrático da função penalidade, como utilizado em (MARTíNEZ, 2000; MATIOLI, 2001), entre outros.

Adotando esta estratégia faremos um ajuste quadrático da função exponencial. Considere $\lambda \in \mathbb{R}_{++}$ fixado e a função $p : \mathbb{R} \to \mathbb{R}$ dada por $p(y) = \lambda e^{\frac{y}{\gamma_k}}$ e $\bar{p} : \mathbb{R} \to \mathbb{R}$ dada por

$$\bar{p}(y) = \begin{cases} \lambda e^{\frac{y}{\gamma_k}} & \text{se } y \leq 0 \\ ay^2 + by + c & \text{se } y > 0\,. \end{cases} \qquad (4.27)$$

O objetivo do ajuste é garantir a suavidade da função, da primeira e segunda derivadas de $\bar{p}$. Denotando $q(y) = ay^2 + by + c$ e $p(y) = \lambda e^{\frac{y}{\gamma_k}}$, às partes quadrática e exponencial, respectivamente, de $\bar{p}$, serão impostas as condições de interpolação:

$$p(0) = q(0),\ p'(0) = q'(0) \text{ e } p''(0) = q''(0).$$

Da condição de interpolação $p(0) = q(0)$ temos $\lambda = c$. A derivada da função $\bar{p}$ é dada por

$$\bar{p}'(y) = \begin{cases} \dfrac{\lambda}{\gamma_k} e^{\frac{y}{\gamma_k}} & \text{se } y \leq 0 \\ 2ay + b & \text{se } y > 0, \end{cases} \qquad (4.28)$$

e usando a condição de interpolação da derivada de $\bar{p}$, $p'(0) = q'(0)$, temos que $b = \frac{\lambda}{\gamma_k}$.

A segunda derivada da função $\bar{p}$ é dada por

$$\bar{p}''(y) = \begin{cases} \dfrac{\lambda}{\gamma_k^2} e^{\frac{y}{\gamma_k}} & \text{se } y \leq 0 \\ 2a & \text{se } y > 0. \end{cases} \tag{4.29}$$

Assim, usando a condição de interpolação na segunda derivada de $\bar{p}$, $p''(0) = q''(0)$ temos que $a = \dfrac{\lambda}{2\gamma_k^2}$.

Logo, q(y) = $\dfrac{\lambda}{2\gamma_k^2} y^2 + \dfrac{\lambda}{\gamma_k} y + \lambda$ e consequentemente,

$$\bar{p}(y) = \begin{cases} \lambda e^{\frac{y}{\gamma_k}} & \text{se } y \leq 0 \\ \dfrac{\lambda}{2\gamma_k^2} y^2 + \dfrac{\lambda}{\gamma_k} y + \lambda & \text{se } y > 0 \end{cases} \tag{4.30}$$

Com a quadrática ajustada, o sistema não linear (4.26) torna-se

$$G^k(x) = \gamma_k(x - x^k) + \nabla F_x(x) + \sum_{i=1}^{m} \bar{p}_i(h(x)) \nabla h_i(x). \tag{4.31}$$

em que $\bar{p}_i$, para $i = 1, \ldots, m$, é dada por (4.30).

No próximo capítulo faremos testes numéricos envolvendo o algoritmo proposto, IALEMe, com e sem o ajuste quadrático da penalidade exponencial.

EXPERIMENTOS NUMÉRICOS

1 I INTRODUÇÃO

Com o intuito de validar a nova metodologia, apresentaremos a seguir resultados numéricos obtidos com a implementação do algoritmo proposto.

O Algoritmo IALEMe, 4.1, foi implementado em MatLab versão 8.1.0.604 (R2013a), executado em um computador MacBook Air com processador 1.4 Intel Core i5 e utilizado na resolução de problemas de equilíbrio geral e problemas de equilíbrio de Nash Generalizado (GNPs). A classe de problemas desiganados GNEPs, sigla derivada de sua denominação em inglês, compõe uma classe particular do problema de equilíbrio geral.

Para iniciar, escreveremos o problema de equilíbrio de Nash Generalizado em termos de um problema de equilíbrio geral. Para tal, descreveremos brevemente o GNEP. Para o estado da arte e uma descrição completa do problema de equilíbrio de Nash Generali- zado pode ser consultado, por exemplo, (FACCHINEI; KANZOW, 2010a; FACCHINEI; KANZOW, 2007) e suas referências.

O termo GNEP designa um jogo não cooperativo com N jogadores no qual cada jogador j controla sua variável de decisão, designada por $x^j \in \mathbb{R}^{n_j}$. A variável de decisão de cada um dos outros jogadores é dada por

$$x^{-j} = (x^1, \ldots, x^{j-1}, x^{j+1}, \ldots, x^N).$$

O vetor $x = (x^j, x^{-j}) = (x^1, \ldots, x^{j-1}, x^j, x^{j+1}, \ldots, x^N) \in X$ denota a variável de decisão de todos os jogadores, em que X $\mathbb{R}^n$ é um conjunto convexo fechado e $n = \sum_{j=1}^{N} n_j$

Formalmente, resolver um GNEP equivale a resolver o problema de

$$\min_{x^j} \theta_j(x^j, x^{-j}) \qquad s.a.: \qquad x^j \in X_j(x^{-j}), \tag{5.1}$$

onde $X_j(x^{-j}) \subseteq \mathbb{R}^{n_j}$ e θ_j, denotam, respectivamente, o conjunto de estratégias e a função de retorno do jogador j para cada $j = 1, \ldots, N$.

Apresentaremos, a seguir, exemplos nos quais o conjunto X tem a forma

$$X = \{x \in \mathbb{R}^n : \omega_i(x) \leq 0, \quad \forall i = 1, \ldots, q\}, \tag{5.2}$$

no qual $\omega_i : \mathbb{R}^n \to \mathbb{R}$ é uma função convexa para cada $i = 1, 2, \ldots, q$.

Para a resolução destes problemas, inicialmente reformulamos o GNEP em termos de um problema de desigualdade variacional, como descrito em (FACCHINEI ANDREAS FISCHER, 2007) e, então, usando o fato que o problema de desigualdade variacional pode ser reformulado como um problema de equilíbrio geral (ver, por exemplo, (IUSEM; SOSA, 2003a)), obtemos um problema de equilíbrio representado por um GNEP cuja função

resultante $f : R^n \times R^n \to R$ e o conjunto $K \to R^n$ têm, respectivamente, a forma

$$f(x,y) = \sum_{j=1}^{N} \left\langle \nabla_{x^j}\theta_j(x), y^j - x^j \right\rangle \; e \; K = X. \tag{5.3}$$

Diante do exposto, o problema da poluição de uma bacia hidrográfica, a ser tratado na próxima seção, é formulado em termos de um GNEP. Contudo, seguindo a discussão acima, este problema pode ser tratado como um problema de equilíbrio. Vale destacar que este é um problema importante e que tem sido objeto de estudos resultando na publicação de diversos artigos. Entre estes, (KRAWCZYK; URYASEV, 2000; MATIOLI; SOSA; YUAN, 2012; SANTOS; SCHEIMBERG, 2011), merecem destaque.

2 | APLICAÇÃO DO ALGORITMO IALEM*E* COM PENALIDADE EXPONENCIAL

Nesta seção apresentaremos exemplos onde aplicamos o algoritmo proposto na resolução de problemas de equilíbrio de Nash Generalizado e problemas de equilíbrio gerais.

O problema da poluição de uma bacia hidrográfica, com 3 jogadores, e sua generalização para N jogadores é o primeiro a ser tratado. O outro, um problema de equilíbrio geral com 5 jogadores e sua generalização para N jogadores. Uma lista completa com estes e outros problemas pode ser encontrada, por exemplo, em (FACCHINEI; KANZOW, 2009).

Exemplo 5.1 *No problema da poluição de uma bacia hidrográfica com três jogadores $j = 1, 2, 3$, estes figuram como possíveis agentes poluidores. O problema consiste em minimizar a função objetivo dada por*

$$\theta_j(x) = c_j \; {x^j}^2 + d_1 x^j(x^1 + x^2 + x^3) - b_i x^j, \qquad j = 1, 2, 3, \tag{5.4}$$

em que $c_1 = c_3 = 0,01$, $c_2 = 0,05$, $d_1 = 0,01$, $b_1 = 2,90$, $b_2 = 2,88$ and $b_3 = 2,85$. As funções ω_i, dadas em (5.2), são definidas por

$$\omega_1(x) = 3,25x^1 + 1,25x^2 + 4,125x^3 - 100$$

e

$$\omega_2(x) = 2,291x^1 + 1,5625x^2 + 2,8125x^3 - 100.$$

Para construirmos o sistema não linear $G^k(x) = 0$, dada em (4.26), será necessário calcular $\nabla F_x(x)$, o qual é dado por

$$\nabla F_x(x) = \begin{bmatrix} 0,04x^1 + 0,01x^2 + 0,01x^3 - 2,90 \\ 0,01x^1 + 0,12x^2 + 0,01x^3 - 2,88 \\ 0,01x^1 + 0,01x^2 + 0,04x^3 - 2,85 \end{bmatrix}.$$

Neste exemplo, os valores iniciais escolhidos foram: $x^0 = [5, 9, 3]^T$, $\lambda^0 = [1, 1.5]^T$ *e* $\gamma_k = 0.4$ para todo k.

Na Tabela 1 a seguir, a primeira coluna representa o nome dos algoritmos utilizados na comparação. A segunda, contém a solução obtida pelos algoritmos comparados e a terceira e quarta colunas são, respectivamente, o número de iterações e o tempo computacional, em segundos, gasto na resolução do problema.

Os algoritmos comparados são:

- *IALEMe: Algoritmo de Lagrangeano Aumentado Extragradiente Exponencial Inexato.*
- *RA: Algoritmo de Relaxação, descrito em (KRAWCZYK; URYASEV, 2000).*
- *MSY: Um Algoritmo Numérico para Encontrar Soluções de um Problema de Equilíbrio de Nash Generalizado, descrito em (MATIOLI; SOSA; YUAN, 2012).*
- *SS: Um Algoritmo de Subgradiente Inexato para Problemas de Equilíbrio, descrito em (SANTOS; SCHEIMBERG, 2011).*
- *EUDA: Algoritmos para o Problema de Equilíbrio Aplicados ao Problema de Equilíbrio de Nash, descritos em (FERREIRA, 2013).*

Algoritmo	Solução obtida	Iterações	tempo
IALEMe	$[21, 144; 16, 028; 2, 7268]^T$	125	0,09
RA	$[21, 149; 16, 028; 2, 722]^T$	20	-
MSY	$[20, 115; 17, 125; 3, 205]^T$	10	1,37
SS	$[21, 144; 16.028\ 2.726]^T$	7	-
EUDA	$[21, 143; 16, 028; 2, 727]^T$	6	-

TABELA 1 – Resultados do Exemplo 5.1

Analisando os resultados da Tabela 1, podemos tirar algumas conclusões. Primeiramente destacamos que algoritmo proposto, IALEMe, utilizou mais iterações para resolver o problema quando comparado com os demais algoritmos. Porém, dispendeu tempo computa- cional menor quando comparado com o algoritmo MSY. O motivo principal do Algoritmo IALEMe utilizar um número maior de iterações tem a ver com o fato da penalidade exponencial ser muito sensível a pequenas alterações nos dados. Isso se deve à presença do parâmetro γ *na expressão do sistema não linear,(4.26), a ser resolvido a cada passo do algoritmo, gerando, em geral, problemas mal condicionados.*

Na sequência executamos o Algortimo IALEMe, aplicando-o ao mesmo problema, porém agora com quadrática ajustada. Os resultados são mostrados na Tabela 2 a seguir, com as colunas sendo as mesmas já descritas na Tabela 1.

Algoritmo	Solução obtida	Iterações	tempo
IALEMe	$[21,144; 16,028; 2,7260]^T$	3	0,04
RA	$[21,149; 16,028; 2,722]^T$	20	-
MSY	$[20,115; 17,125; 3,205]^T$	10	1,37
SS	$[21,144; 16,028; 2,726]^T$	7	-
EUDA	$[21,143; 16,028; 2,727]^T$	6	-

TABELA 2 – Resultados do Exemplo 5.1 com quadrática ajustada

Observe agora, nos resultados da Tabela 2, que houve grande melhoria no desempenho do Algoritmo IALEMe.

O ajuste quadrático consegue contornar o problema do mau condicionamento, ou seja, o algoritmo ficou menos dependente do valor de γ na expressão do sistema não linear em (4.26). Ainda, podemos observar que, o algoritmo apresenta desempenho melhor em relação aos demais com os quais foi comparado.

Exemplo 5.2 *Neste exemplo, consideramos a generalização do problema da poluição de uma bacia hidrográfica para N jogadores ao invés de 3. Os testes são executados variando-se o valor de N. Neste caso, a função* θ *de cada jogador é dada por*

$$\theta_j(x) = u_j(x^j)^2 + \beta x^j \left(\sum_{i=1}^{N} x^i\right) - v_j x^j \quad \forall\, j = 1, \ldots, N,$$

onde $\beta = 0,01$ *e* $u_j, v_j \in \mathbb{R}$ *para cada* $j = 1, \ldots, N$. *Com esta generalização temos m restrições, as quais representam as constantes de emissão de poluição, definidas como* $Ax \leq b$, *onde* $A \in \mathbb{R}^{m \times N}$ *e* $b \in \mathbb{R}^m$. *O vetor de multiplicadores de Lagrange* λ^0, *a matriz A, o vetor b e os escalares* u_j *e* v_j *são gerados randomicamente. O ponto inicial foi escolhido como sendo* $x^0 = [1, 1, \ldots, 1]^T$. *A função f associada ao problema de equilíbrio nesta configuração é dada por* (5.3) *e o conjunto K é dado por* $K = X = \{x \in \mathbb{R}^N : Ax \leq b\}$.

Na Tabela 3 a seguir, a primeira coluna representa o nome do algoritmo, a segunda o número de jogadores, a terceira o número de restrições, a quarta o número de iterações executadas e a última *coluna representa o tempo, em segundos, gasto para resolver o problema em questão.*

Constatamos que, para este problema, as respostas são melhores quando o parâmetro γ é pequeno. Por isso, nos testes consideramos $\gamma_k = 1e-5$. *Além disso, para problemas maiores (acima de 1000 jogadores), o método não consegue obter uma solução*

para o problema devido ao mau condicionamento gerado pela penalidade exponencial. Esta dificul- dade é contornada com a utilização da penalidade exponencial com quadrática ajustada. Apresentaremos os resultados desse procedimento nos testes da próxima seção.

Algoritmo	*N*	*m*	Iterações	Tempo
IALEMe	10	5	1	0,03
IALEMe	30	20	2	0,03
IALEMe	50	10	2	0,04
IALEMe	100	10	2	0,4
IALEMe	200	20	2	0,12

TABELA 3 – Resultados do Exemplo 5.2

Como os dados desse exemplo foram gerados randomicamente não seria conveniente realizar comparações. De qualquer forma, este mesmo exemplo foi apresentado em (FERREIRA, 2013; MATIOLI; SOSA; YUAN, 2012) e os resultados são parecidos. Mas ainda persiste o problema do mau condicionamento gerado pela penalidade exponencial e só conseguimos resolver problemas com dimensões não muito altas, com até 200 jogadores e 20 restrições, como mostrado na tabela acima. Por isso, esse mesmo exemplo foi executado usando a penalidade com quadrática ajustada e com os mesmos dados iniciais, ou seja, geramos randomicamente os dados e salvamos em disco para executar e comparar as duas abordagens, com e sem quadrática ajustada. Os resultados são mostrados na Tabela 4 a seguir.

Algoritmo	*N*	*m*	Iterações	Tempo
IALEMe	10	5	1	0,03
IALEMe	30	20	2	0,03
IALEMe	50	10	3	0,04
IALEMe	100	10	4	0,06
IALEMe	200	20	2	0,11
IALEMe	1000	20	7	70,49
IALEMe	1000	100	9	289,34

TABELA 4 – Resultados do Exemplo 5.2 com quadrática ajustada

Observe que, com o ajuste quadrático, foi possível resolver problemas maiores com até 1000 jogadores e 100 restrições, o que não ocorria sem o ajuste quadrático.

Exemplo 5.3 *Neste exemplo, consideramos um problema de equilíbrio quadrático EP(f_i, K), para i =* 1*,* 2*, introduzido em em (SANTOS; SCHEIMBERG, 2011) e também em*

(TRAN; DUNG; NGUYEN, 2008), onde

$$K = \{x \in \mathbb{R}^5 : \sum_{j=1}^{5} x^j \geq -1, -5 \leq x^j \; 5, j \leq 1, \ldots, 5\}$$

e

$$f_i(x, y) = \langle P_i x + Qy + q, y - x \rangle, \qquad i = 1, 2, \qquad (5.5)$$

na qual

$$P_1 = \begin{bmatrix} 3,1 & 2 & 0 & 0 & 0 \\ 2 & 3,6 & 0 & 0 & 0 \\ 0 & 0 & 3,5 & 2 & 0 \\ 0 & 0 & 2 & 3,3 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 3,1 & 2 & 0 & 0 & 0 \\ 2 & 3,6 & 0 & 0 & 0 \\ 0 & 0 & 3,5 & 2 & 0 \\ 0 & 0 & 2 & 3,3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix},$$

$$Q = \begin{bmatrix} 1,6 & 1 & 0 & 0 & 0 \\ 1 & 1,6 & 0 & 0 & 0 \\ 0 & 0 & 1,5 & 1 & 0 \\ 0 & 0 & 1 & 1,5 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \quad e \;\; q = \begin{bmatrix} 1 \\ -2 \\ -1 \\ 2 \\ -1 \end{bmatrix}.$$

Os valores iniciais considerados para este exemplo foram:

$$x^0 = [1, 3, 1, 1, 2]T, \qquad \lambda^0 = [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]^T \qquad e\ \gamma_k = 1e - 5.$$

Observe que a restrição $-5 \leq x^j \leq 5$ *é equivalente a* $x^j - 5 \leq 0$ *e* $-x^j - 5 \leq 0$*, para todo* $j = 1, \ldots, 5$*. Os resultados numéricos são mostrados nas Tabelas 5 and 6 a seguir. Os algoritmos comparados são:*

- *IALEMe: Algoritmo de Lagrangeano Aumentado Extragradiente Exponencial Inexato.*
- *EA: Algoritmo Extragradiente Estendido para o Problema de Equilíbrio, descrito em (TRAN; DUNG; NGUYEN, 2008).*
- *SS: Um Algoritmo de Subgradiente Inexato para Problemas de Equilíbrio, descrito em (SANTOS; SCHEIMBERG, 2011).*

Algoritmo	Solução obtida	Iterações	Tempo
IALEMe	$[-0,725; 0,803; 0,720; -0,867; 0,250]^T$	1	0,04
EA-a	$[-0,726; 0,804; 0,720; -0,866; 0,250]^T$	10	-
SS	-	10	0,0006

TABELA 5 – Resultados do Exemplo 5.3 associado com $EP(f_1; K)$

Algoritmo	Solução obtida	Iterações	Tempo
IALEMe	$[-0,725; 0,803; 0,720; -0,867; 0,200]^T$	1	0,04
EA-b	$[-0,726; 0,803; 0,718; -0,864; 0,200]^T$	21	-
SS	-	10	0,0006

TABELA 6 – Resultados do Exemplo 5.3 associado com $EP\,(f_2; K)$

Para este exemplo, os autores (TRAN; DUNG; NGUYEN, 2008) não exibem o tempo computacional para o Algoritmo EA, para nenhum dos dois casos a e b. Se compararmos somente o número de iterações, o Algoritmo IALEMe foi superior. Por outro lado, os autores (SANTOS; SCHEIMBERG, 2011) não exibem a solução para o Algoritmo SS. Comparando então, o número de iterações, o Algoritmo IALEMe foi superior ao Algoritmo SS, porém o inverso ocorreu em relação ao tempo computacional.

Exemplo 5.4 *Neste exemplo, consideramos a generalização do problema anterior para N jogadores ao invés de 5. Desta forma, o conjunto K é dado por*

$$K = \{\, x \in \mathbb{R}^N : \sum_{j=1}^{N} x^j \geq -1, -5 \leq x^j \quad 5,\ \ j \leq 1, \ldots, N\}.$$

O número de restrições é $m = 2N + 1$, as matrizes P_i, para $i = 1, 2$, Q e o vetor q são gerados randomicamente. Os resultados numéricos são mostrados na Table 7 a seguir, com o parâmetro $\gamma_k = 1e - 5$ para toda iteração k.

Algoritmo	N	m	Iterações	Tempo
IALEMe	10	21	1	0,04
IALEMe	20	41	1	0,04
IALEMe	50	101	1	0,06
IALEMe	100	201	1	0,09
IALEMe	150	301	1	0,17

TABELA 7 – Resultados do Exemplo 5.4

Aqui também, os dados foram gerados randomicamente, e por isso as comparações com outros métodos não seriam convenientes. Mas, como ocorreu nos exemplos anteriores, quando utilizamos a penalidade exponencial sem o ajuste quadrático, só conseguimos resolver o problema com até 200 jogadores e 20 restrições.

Na próxima tabela, listamos os resultados alcançados pelo Algoritmo IALEMe, mas agora com quadrática ajustada. Neste caso, para contrastarmos as duas abordagens, manti- vemos o valor de $\gamma_k = 1e - 5$ para os testes com até 200 jogadores e 20 restrições, que foram aqueles que o Algoritmo IALEMe, sem o ajuste quadrático, conseguiu resolver.

Para os demais testes, ou seja, acima de 200 jogadores, o Algoritmo IALEMe se comportou melhor quando γ_k é maior. Por isso adotamos o valor $\gamma_k = 5$*, para todas as iterações k (apesar de resolver também com outros valores para* γ*, inclusive o valor* $\gamma_k = 1e-5$*, para todas as iterações k mas com desempenho um pouco pior).*

Algoritmo	*N*	*m*	Iterações	Tempo
IALEMe	10	21	1	0,04
IALEMe	20	41	1	0,04
IALEMe	50	101	1	0,06
IALEMe	100	201	1	0,08
IALEMe	150	301	1	0,09
IALEMe	300	601	3	0,91
IALEMe	500	1001	3	1,5
IALEMe	1000	2001	3	11,2
IALEMe	2000	4001	3	83,8
IALEMe	3000	6001	3	291
IALEMe	4000	8001	3	683

TABELA 8 – Resultados do Exemplo 5.4 com quadrática ajustada

CONSIDERAÇÕES FINAIS

Inicialmente procedemos a uma revisão bibliográfica dos tópicos de otimização necessários para o desenvolvimento da proposta desta tese. Destacamos como conceitos principais aqueles relacionados aos métodos de lagrangeano aumentado e ao problema de equilíbrio geral.

Em um primeiro momento pretendíamos avançar no sentido de estender os algoritmos propostos em (NASRI, 2010) e (FERREIRA, 2013) para resolver o problema de equilíbrio geral. Cumprimos este objetivo propondo um novo algoritmo de lagrangeano aumentado com penalidade exponencial, em substituição à penalidade clássica de Rockafellar que estes autores utilizaram em seus trabalhos. Com esta abordagem introduzimos o Algoritmo IALEMe, o qual sintetizou a principal contribuição deste trabalho.

Na sequência, procuramos refazer a teoria geral em torno do novo algoritmo. Mostramos que a sequência de multiplicadores (dual), gerada pelo Algoritmo IALEMe, coincide com a sequência gerada pelo Algoritmo IPPEMIe com distância de Bregman, assim como foi realizado em (NASRI, 2010).

Sobre a convergência conjecturamos que o algoritmo é convergente. Conseguimos mostrar uma espécie de convergência local impondo uma condição sobre o vetor de multiplicadores de Lagrange.

Em um segundo momento, com o objetivo de contornar os problemas de mau condicionamento decorrentes do uso da penalidade exponencial, realizamos um ajuste quadrático desta. Realizamos testes numéricos empregando a penalidade exponencial e a também utilizando esta com quadrática ajustada.

O algoritmo proposto foi implementado em MatLab e aplicado na resolução de problemas de equilíbrio e problemas de equilíbrio de Nash Generalizados, já conhecidos na literatura. Comparamos os resultados alcançados pelo IALEMe com outros trabalhos já publicados.

Os testes numéricos mostraram que, utilizando o ajuste quadrático na penalidade exponencial, o IALEMe teve desempenho superior aos algoritmos com os quais foi compa- rado. Por outro lado, quando utilizamos a penalidade exponencial sem o ajuste quadrático, o algoritmo pode gerar subproblemas mal condicionados, ou seja, dependente do parâmetro de penalidade. No entanto, a penalidade exponencial foi importante para mostrar os principais resultados da tese.

Como sugestões para trabalhos futuros, destacamos:

- Apresentar novos algoritmos com penalidades diferentes da exponencial, tal como a logarítmica quadrática, com a qual tem-se obtido algum sucesso em trabalhos envolvendo o método de lagrangeano aumenado para o problema de otimização. Acreditamos que o mesmo possa ocorrer com o problema de equilíbrio geral, desde que o problema de otimização pode ser reformulado como

um caso particular deste.

- Realizar testes numéricos com outros problemas de equilíbrio geral, incluindo casos particulares destes, como por exemplo, desigualdade variacional.
- Pesquisar estratégias diferentes para a escolha do parâmetro de penalidade γ tais como aquelas desenvolvidas em (BIRGIN; MARTíNEZ, 2012).

REFERÊNCIAS

BERTSEKAS, D. P. *Multiplier Methods: A Survey*. [S.l.: s.n.], 1976.

BIANCHI, M.; PINI, R. Coercivity conditions for equilibrium problems. *Journal of Optimization Theory and Applications*, v. 124, n. 1, p. 79–92, 2005.

BIGI, G. et al. Existence and solution methods for equilibria. *European Journal of Operational Research*, v. 227, n. 1, p. 1–11, 2013.

BIRGIN, E.; MARTíNEZ, J. M. Augmented lagrangian method with nonmonotone penalty parameters for constrained optimization. *Computational Optimization and Applications*, v. 51, n. 3, p. 941–965, 2012.

BLUM, E.; OETTLI, W. From optimization and variational inequalities to equilibrium problems. *The Mathematics Student*, v. 63, p. 123–145, 1994.

BURACHIK, G. K. R. On a generalized proximal point method for solving equilibrium problems in banach spaces. *Nonlinear Analysis: Theory, Methods & Applications*, v. 75, n. 18, p. 6456–6464, 2012.

EUSTÁQUIO, R. G. *Condições de otimalidade e de qualificação para problemas de programação não linear*. Dissertação (Mestrado), 2007.

FACCHINEI ANDREAS FISCHER, V. P. F. On generalized nash games and variational inequalities. *Operations Research Letters*, v. 35, n. 2, p. 159–164, 2007.

FACCHINEI, F.; KANZOW, C. Generalized nash equilibrium problems. *4OR*, Springer-Verlag, v. 5, n. 3, p. 173–210, 2007.

FACCHINEI, F.; KANZOW, C. Penalty methods for the solution of generalized nash equilibrium problems (with complete test problems). *Sapienza University of Rome*, 2009.

FACCHINEI, F.; KANZOW, C. Generalized nash equilibrium problems. *Annals of Operations Research*, v. 175, n. 1, p. 177–211, 2010.

FACCHINEI, F.; KANZOW, C. Penalty methods for the solution of generalized nash equilibrium problems. *SIAM Journal on Optimization*, v. 20, n. 5, p. 2228–2253, 2010.

FACCHINEI, J.-S. P. F. *Finite-Dimensional Variational Inequalities and Complementarity Problems*. 1. ed. [S.l.: s.n.], 2003.

FERREIRA, R. G. *Algoritmos para o problema de Equilíbrio de Nash*. Tese (Doutorado), 2013.

FLåM, S. D.; ANTIPIN, A. S. Equilibrium programming using proximal-like algorithms. *Mathematical Programming*, v. 78, n. 1, p. 29–41, 1996.

GRIVA, I.; NASH, S.; SOFER, A. *Linear and Nonlinear Optimization: Second Edition*. [S.l.]: Society for Industrial and Applied Mathematics (SIAM, 3600 Market Street, Floor 6, Philadelphia, PA 19104), 2009.

HESTENES, M. R. Multiplier and gradient methods. *Journal of Optimization Theory and Applications*, v. 4, n. 5, p. 303–320, 1969.

HIRIART-URRUTY, J.-B.; Lemaréchal, C. *Convex Analysis and Minimization Algorithms*. [S.l.]: Springer Verlag, Heidelberg, 1996. Two volumes - 2nd printing.

IUSEM, A.; KASSAY, G.; SOSA, W. On certain conditions for the existence of solutions of equilibrium problems. *Mathematical Programming*, v. 116, n. 1-2, p. 259–273, 2009.

IUSEM, A. N. *Métodos de Ponto Proximal em Otimização*. [S.l.]: 20o Colóquio Brasileiro de Matemática - IMPA, 1995.

IUSEM, A. N.; NASRI, M. Inexact proximal point methods for equilibrium problems in banach spaces. *Numerical Functional Analysis and Optimization*, v. 28, n. 11-12, p. 1279–1308, 2007.

IUSEM, A. N.; SOSA, W. Iterative algorithms for equilibrium problems. *Optimization*, v. 52, n. 3, p. 301–316, 2003.

IUSEM, A. N.; SOSA, W. New existence results for equilibrium problems. *Nonlinear Analysis: Theory, Methods & Applications*, v. 52, n. 2, p. 621–635, 2003.

IUSEM, A. N.; TEBOULLE, M. Convergence rate of nonquadratic proximal point methods for convex and linear programming. *hMathematics of Operations Reserc*, n. 8, p. 657 – 677, 1995.

KARMARKAR, N. K. A new polynomial-time algorithm for linear programming. p. 302–311, 1984.

KRAWCZYK, J. B.; URYASEV, S. Relaxation algorithms to find nash equilibria with economic applications. *Environmental Modeling & Assessment*, v. 5, n. 1, p. 63–73, 2000.

LUENBERGER, Y. Y. D. G. *Linear and Nonlinear Programming*. [S.l.]: Springer US, 2008.

MARTíNEZ, J. M. Box-quacan and the implementation of augmented lagrangian algorithms for minimization with inequality constraints. *Computational & Applied Mathematics*, v. 19, p. 31–56, 2000.

MASTROENI, G. On auxiliary principle for equilibrium problems. Publicatione del Dipartimento di Mathematica Dell Universita di Pisa, v. 3, p. 1244–1258, 2000.

MATIOLI, L. C. *Uma nova metodologia para construção de funções de penalização para algoritmos de lagrangeano aumentado*. Tese (Doutorado), 2001.

MATIOLI, L. C.; SOSA, W.; YUAN, J. A numerical algorithm for finding solutions of a generalized nash equilibrium problem. *Computational Optimization and Applications*, v. 52, n. 1, p. 281–292, 2012.

MINOUX, M. *Mathematical programming: theory and algorithms*. [S.l.]: Wiley, 1986. (Wiley-Interscience series in discrete mathematics and optimization).

MUU, L.; OETTLI, W. Convergence of an adaptive penalty scheme for finding constrained equilibria. *Nonlinear Analysis: Theory, Methods & Applications*, v. 18, n. 12, p. 1159–1166, 1992.

MUU, L. D.; QUOC, T. D. Regularization algorithms for solving monotone ky fan inequalities with application to a nash-cournot equilibrium model. *Journal of Optimization Theory and Applications*, v. 142, n. 1, p. 185–204, 2009.

NASRI, A. N. I. . M. Augmented lagrangian methods for equilibrium problems. *RAIRO, Recherche Opérationnelle*, v. 44, p. 5–26, 2010.

NGUYEN, T.; STRODIOT, J.-J.; NGUYEN, V. A bundle method for solving equilibrium problems. *Mathematical programming*, v. 116, n. 1-2, p. 529–552, 2009.

NGUYEN, T. T. V.; STRODIOT, J.-J. et al. The interior proximal extragradient method for solving equilibrium problems. *Journal of Global Optimization*, v. 44, n. 2, p. 175–192, 2009.

POWELL, M. A method for nonlinear constraints in minimization problems. In: FLETCHER, R. (Ed.). *Optimization*. [S.l.]: Academic Press, 1969. p. 283–298.

QUOC, T. D.; MUU, L. D. Implementable quadratic regularization methods for solving pseudomonotone equilibrium problems. *East-West Journal of Mathematics*, v. 6, n. 2, 2004.

ROCKAFELLAR, R. A dual approach to solving nonlinear programming problems by unconstrained optimization. *Matematical Programming*, v. 5, p. 354–373, 1973.

ROCKAFELLAR, R. T. Augmented lagrange multiplier functions and duality in nonconvex programming. *SIAM Journal on Control*, v. 12, n. 2, p. 268–285, 1974.

SANTOS, J. M. M. e S. A. *Métodos Computacionais de Otimização*. [S.l.]: 20o Colóquio Brasileiro de Matemática - IMPA, 1998.

SANTOS, J. M. M. e S. A. *Métodos Computacionais de Otimização*. [S.l.]: Departamento de Matemática Aplicada - IMECC - UNICAMP, 1999.

SANTOS, P.; SCHEIMBERG, S. An inexact subgradient algorithm for equilibrium problems. *Computational & Applied Mathematics*, v. 30, p. 91–107, 2011.

SOLODOV, A. I. e M. *Otimização v. 1: Condições de Otimalidade, Elementos de Análise Convexa e de Dualidade*. [S.l.]: Instituto Brasileiro de Matemática Pura e Aplicada - IMPA, 2005.

TRAN, D. Q.; DUNG, M. L.; NGUYEN, V. H. Extragradient algorithms extended to equilibrium problems. *Optimization*, v. 57, n. 6, p. 749–776, 2008.

ADILSON DA SILVEIRA - Professor efetivo do Departamento de Matemática na Universidade Tecnológica Federal do Paraná - UTFPR– câmpus Pato Branco. Formação acadêmica em licenciado em Ciências/ Matemática pela UTFPR nesse mesmo câmpus, possui Mestrado e Doutorado em Métodos Numéricos em Engenharia pela Universidade Federal do Paraná - UFPR, na linha de pesquisa de Análise Numérica e Otimização.

LAGRANGEANO AUMENTADO EXPONENCIAL APLICADO AO PROBLEMA DE EQUILÍBRIO

www.atenaeditora.com.br

contato@atenaeditora.com.br

@atenaeditora

www.facebook.com/atenaeditora.com.br

Ano 2023

LAGRANGEANO AUMENTADO EXPONENCIAL APLICADO AO PROBLEMA DE EQUILÍBRIO

www.atenaeditora.com.br

contato@atenaeditora.com.br

@atenaeditora

www.facebook.com/atenaeditora.com.br

Ano 2023